5-2-2018

The Cognition of Distrust: Pathways To Attitudes

Justin Mahalak
University of Connecticut - Storrs, jgmahalak@gmail.com

The Cognition of Distrust: Pathways To Attitudes

Justin G. Mahalak, Ph.D.

University of Connecticut, 2018

Six experiments examined the effects of diffuse distrust on cognition, in order to elucidate the link between the immediate cognitive changes associated with the state and responses to persuasive messages. Five of the present studies addressed two specific mechanisms that appear to account for established non-routine effects – increased contrast sensitivity and cognitive flexibility. Experiment 1 used an imposter detection task (IDT) to induce distrust. It presented participants with 30 statements, to rate the likelihood of each on a 100-point scale. The second half consisted of statements that agreed with, contradicted, or were unrelated to those in the first half. Impostor detection did not have the predicted contrast-sensitizing effect on these plausibility judgments, but values-based medical mistrust was associated with higher judged plausibility of factually-false statements that did not contradict previously-presented statements. Experiment 2 induced distrust via economic deception game (EDG). Four characters, including a medical doctor and a quack, were presented either together or alone. The predicted effect, that doctor and charlatan would be rated less stereotypically by distrusting participants when presented alone, and more when presented together, was not found. Experiment 3 did not find the predicted result that the EDG would make participants agree less with strong arguments for senior comprehensive exams, and more with an anti-exam argument. In Experiment 4, the IDT did not increase completed word stems. In Experiment 5, after the EDG, participants rated the compatibility of statement pairs, and provided explanations of how to make them more compatible. Participants in the distrust condition who believed their partner was real used more words to reconcile the statements, as predicted. Experiment 6 used the EDG and presented participants with a pro-tire-rotation essay, and an anti-tire-rotation essay designed to imitate the rhetorical style of anti-vaccination websites. Participants in the distrust condition were generally more skeptical of the anti-rotation essay, but if they scored highly on medical mistrust, they found it more convincing. The relevance of basic cognitive processes to explicit

attitudes is discussed, and it is concluded that existing beliefs may have important implications for how people respond to deception.

The Cognition Of Distrust: Pathways To Attitudes

Justin G. Mahalak

B.S., Eastern Michigan University, **2007**

M.A., New York University, **2010**

A Dissertation

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Doctor of Philosophy

at the

University of Connecticut

2018

APPROVAL PAGE

Doctor of Philosophy Dissertation

The Cognition Of Distrust: Pathways To Attitudes

Presented by

Justin G. Mahalak, B.S., M.A.

Major Advisor _____
Kerry L. Marsh, Ph.D.

Associate Advisor _____
Blair T. Johnson, Ph.D.

Associate Advisor _____
Dev Dalal, Ph.D.

University of Connecticut
2018

Acknowledgments

Ashley – For endless tolerance of my ramblings, for unwavering personal support, and for help with MS Office. I love you.

Kerry – For her patience, her advice, and for giving me this opportunity in the first place, years ago after I sent out applications for what I had decided would be the last time.

My Parents – For always prioritizing my education. And for their optimism, even when I don't want to hear it.

Alphonso and Ricky – for the lab work, always done faster than I could find more for them.

Leon – May you always be worthy of trust.

Table of Contents

Chapter 1

What Is Distrust?

In order for any human society to function effectively, interpersonal and institutional trust must be present to some healthy degree. A thriving economic system relies on the ability of people to be confident that they will get what they were promised. Public safety and peace entirely depend on whether the law and those who enforce it act with integrity and fairness. Adequately dependable news and educational institutions are necessary for keeping people cognizant of what is going on in the world, and how to make the right decisions for their lives. Health messages in particular, if they are to be of any use at all in fighting disease, must be trustworthy. With public trust in institutions eroding, all of these public goods are increasingly at risk (Twenge, Campbell, & Carter, 2014; Wood & Berg, 2011). We should therefore try to understand the psychology of distrust, to understand why it occurs, how it changes the way our mental processes work, and in what ways it can influence a person's decisions. This investigation purports to advance our understanding of how the cognitive effects of distrust influence attitudes.

Put in very simple terms, distrust is what we experience when we feel deceived or vulnerable to deception. Uncertainty is a necessary element of distrust, but some element of human intention seems to also be necessary to activate the characteristic cognitive processes that researchers have identified in laboratory studies of distrust (Mayer & Mussweiler, 2011; Posten & Mussweiler, 2013; Schul, Mayo, & Burnstein, 2004, 2008). So, in order to understand distrust, we first need to understand how people respond to the unpleasant predicament of experiencing a loss of control and certainty about their surroundings.

Distrust and Uncertainty/Control Motivation

Uncertainty and loss of control shake us out of our default ways of operating and force us to engage more critically with our surroundings. Participants who experience a loss of control exhibit a tendency to evaluate information more carefully, to actively seek diagnostic information, and ultimately to respond with more accuracy in laboratory tests. They also display a more negative mood (Pittman & D'Agostino, 1985, 1989; Pittman & Pittman, 1980).

Pittman and Pittman (1980) found that giving participants arbitrary, essentially random feedback on the correctness or incorrectness of their choices in a guessing game left participants with a more negative mood, yet improved their performance on an anagram task in comparison to participants who received no feedback at all on their choices. Using the same paradigm, Pittman and D'Agostino (1989) found that participants receiving random feedback inferred traits based on behaviors more quickly, spent more time studying textual information, and responded more accurately in a recall test based on said textual information.

Studies of this kind involve an experimenter providing the participant with unreliable information, thus being rather unreliable, though not necessarily malicious – that is, they cannot be *trusted*, but this does not mean they are *distrusted*. What they do not address, however, is how people deal with uncertainty that is understood to be purely chance-based, such as that generated by a machine or by the emergent behavior of many people (e.g. fluctuations in stock prices), *versus* when that uncertainty is believed to come from the intentional actions of another person.

In order to parse this difference between chance-based and deception-based uncertainty, Schul, Mayo, Burnstein, and Yahalom (2007) presented participants with a task in which they had the opportunity to win additional money by correctly guessing the content of a matchbox. All participants were told that the outside of the box would match the inside in ⅔ of cases.

Participants in the *chance* condition pulled the matchbox out of a bag, and participants in the *deception* condition had the box handed to them by a partner who stood to profit from their incorrect guesses. The researchers found that participants whose uncertainty was based on random chance tended to adopt a *statistical* approach - they based their guess on the sticker on the outside of the matchbox, accepting that they would be incorrect about ⅓ of the time. On the other hand, participants experiencing deception-based uncertainty adopted a *clinical* approach - they tried to construct a narrative around their partner's actions and predict the contents of each box according to their hypothesis about his or her deception plan. These participants performed considerably worse.

Based on these findings, it appears that when uncertainty is believed to be chance-based, people will adopt a strategy aimed at maximizing successful outcomes, accepting that some error will be inevitable. However, when the uncertainty is attributed to a distrusted other, people will try to eliminate the uncertainty - either by avoiding the distrusted entity altogether, trying to read their mind, as Schul et al.'s participants did, or by relying on socially-constructed incentive structures to restrict (and therefore make more predictable) the other's choices. Uncertainty, therefore, is a necessary but insufficient element of distrust. In order for distrust to be experienced, uncertainty must be accompanied by a degree of vulnerability to the intentional actions of another.

This helps to explain why some recent studies have found that dispositional measures of trust and distrust are addressing different constructs. Conchie, Taylor, and Charlton (2011) found that trust and distrust in leaders were both influenced by the perceived integrity of a given leader, but that the perceived ability of the leader only enhanced trust, having no effect on a person's distrust of the leader. They also found that the perceived benevolence of a leader had very little

influence over how much the leader was trusted, but that perceived malevolence strongly predicted distrust. Ou and Sia (2010) have also demonstrated that trust and distrust have distinct cognitions, antecedents, and consequences, behaving more like separate constructs than locations on a continuum. Similarly, major questionnaire measures of trust (e.g., Shea et al., 2008; Yamagishi et al., 2015; Yamagishi & Yamagishi, 1994) are often split into two factors – one that addresses the need for caution or the perceived competence of others, concepts that deal with basic human tendencies toward error, and one that addresses the goodness or values of others, indicators that they will intentionally behave in ways that are detrimental to oneself.

Focused and Diffuse Distrust

Schul and Peri (2015) have suggested that there is another way that distrust can be conceptually split. They contend that there are two strategic necessities that come into play when our distrust is *focused* on a particular entity – "unmasking the other" and "masking the self."

Focused distrust. In our attempts to unmask the other, we engage in more broad interpretations of their statements and actions, and try to detect inconsistencies, which could reveal deception attempts. We engage in non-routine cognitions in response to a situation in which things may not be what they seem – where it is possible that some messages are decoys, falsehoods, or other means to manipulate us into being taken advantage of. With this strategy, we are trying to answer the question, "what is this really about?" Does this salesperson mean that this car is reliable in the way that I understand the term, or is it just a reliably-selling product for her? Perhaps she is only emphasizing its reliability in order to draw my attention away from its flaws. Perhaps she is lying about the car's reliability outright – is there any available information to contradict her claim? We entertain a wider range of interpretations of the information available, and we look for ways to test these interpretations.

Masking the self is also necessary during such an encounter, because when we are wary that someone may try to deceive or otherwise manipulate us, we feel the need to conceal our own thoughts and motives, in order to make ourselves less vulnerable to manipulation. With this strategy, we tend to become more rigid, reluctant to show signs of weakness to the distrusted other, and relying more on our own "gut" feelings to make decisions, resisting making the decisions we believe they want us to make. If we distrust the car salesperson, this may take the form of maintaining a "poker face" while she pitches the car, or looking for problems to point out with the car when we are actually enamored with it. If we reveal any weakness in our mental armor to the distrusted other, we run the risk of losing our autonomy altogether, so even when we find her persuasive, we try to conceal it. In this way, distrusting others makes us behave in less trustworthy ways, ourselves.

Diffuse distrust. When the encounter that prompted a distrust response has ended, and we are no longer directly interacting with the distrusted other, we are left dealing with the residual activation of these distrust-oriented cognitive processes for a time. Strategically, the need to "mask the self" is no longer present, but we continue to "unmask the other," internally. The creative, critical, way of thinking remains at work until we settle back into our normal cognitive habits.

Studies that observe the effects of distrust in this diffuse form have found an increase in what is often referred to as "non-routine" cognition – the influence of inadmissible information is weakened for participants in mock juries (Fein, McCloskey, & Tomlinson, 1997), the halo effect is reduced (Schul, Burnstein, & Bardi, 1996), counterintuitive puzzles are solved more easily, whereas straightforward ones become more difficult (Schul et al., 2008), creativity of responses increases (Mayer & Mussweiler, 2011), and stereotyping is reduced (Posten & Mussweiler,

2013). This shift away from our personal biases in judgment tends to make us *more* likely to take another person's advice, which is the opposite of what happens if that person is subject to *focused* distrust (Schul & Peri, 2015).

This shift is not necessarily advantageous. Despite the fact that over-reliance on personal biases can often lead us astray, if a bias is, say, in favor of legitimate scientific or professional authorities, it can be a helpful shortcut or heuristic. If the psychology of diffuse distrust is used to undermine the legitimacy of critical institutions and make half-baked alternatives sound more promising, for instance, the potential for harm is great.

Chapter 2

Why Do People Distrust?

If distrust is a departure from routine modes of cognition, then it must also be a response to some sort of environmental cue. Some of these cues have been documented in the psychological literature.

Characteristics of the message or messenger.

Message characteristics. Sometimes information of shaky validity can promote a state of diffuse distrust. One obvious way this can happen is when a message or messenger provides information the recipient knows to be false. In laboratory experiments, distrust toward a particular set of individuals has been provoked by consistently attributing false statements to those individuals (Schul, Mayo, & Burnstein, 2004).

It also may be possible to induce a state of distrust in a less direct way--without explicit experience with false information. For example, the semantic content of a message can cue associations the recipient has previously associated with distrust, putting its recipient into a more distrusting state of mind. Various experiments have used semantic primes (distrust-related words) embedded in scrambled sentence tasks (Friesen & Sinclair, 2011; Mayer & Mussweiler, 2011; Posten & Mussweiler, 2013), and lexical decision tasks (Mayer & Mussweiler, 2011; Posten, Ockenfels, & Mussweiler, 2014; Posten & Mussweiler, 2013) to promote distrust. This suggests that even the words of a message may subtly induce a state of distrust in their recipient.

Communicator characteristics. Aside from the message itself, the messenger's features might also act as cues to distrust. They may convey a certain facial expression, for example. Schul, Mayo, and Burnstein (2008) promoted distrust in laboratory participants simply by having them fixate on and form an impression of an illustrated human face that was designed to look

particularly untrustworthy – a thin face with a goatee and narrowed eyes (based on empirical findings about perceived suspiciousness and facial expressions, see Akehurst, Köhnken, Vrij, & Bull, 1996). So specific facial features and expressions may make a person appear to be less trustworthy, and as a consequence, also promote a distrusting state in perceivers who were exposed to an image of such a person. This effect may require a certain minimum amount of time and attention to work. A previous study had used similarly-designed faces that were presented for much shorter periods of time, and did not find evidence that the facial expressions promoted distrust (Schul et al., 2004).

Social Context

Surveilling and being surveilled by others. Research on the nature of trust and distrust within organizations seems to suggest that surveillance promotes a state of distrust in both the watcher and the watched.

Cialdini notes in his Triple Tumor model of organizational dishonesty that surveillance can have a poisonous effect on the climate of an organization (Cialdini, 1996). Through basic cognitive dissonance processes, supervisors who are surveilling their subordinates will come to attribute honest employee behavior to their surveillance practices rather than to the trustworthiness of the employees. This leads to a situation in which surveillance justifies itself, creating a feedback loop in which investment in surveillance equipment and services promotes further investment in the same, leaching resources that organizations could use in more productive ways.

Additionally, placing people under evaluative scrutiny tends to promote feelings of personal persecution and paranoid cognition in general – when we feel we are being watched by others, we tend to overestimate how much attention is being directed at ourselves, leading us in

turn to overestimate how much of our supervisors' behaviors are directed at ourselves, creating a pattern of heightening suspicion of the motives of others (Kramer, 1994). This pattern of increased cognitive elaboration is conceptually quite similar to that observed in studies of the diffuse state of distrust (Schul et al., 2003;2008; Posten & Mussweiler 2013; Mayer & Mussweiler, 2011).

If we are already treating someone as if they are going to behave in an untrustworthy way, and they then behave in a way that is entirely trustworthy, we are still likely to continue to distrust them. This is true not only because of the dissonance-based mechanisms described in the triple tumor model, but also, distrust drives us to discount information from the distrusted source, and to avoid contact – both of which make it unlikely that we will encounter credible information to effectively challenge our distrust of that entity (Kramer, 1999).

Awareness of interests in conflict with one's own. Any two people who have mutually conflicting interests will have difficulty trusting one another, so even the awareness of the fact that another person has an incentive to deceive us may lead us to distrust them. One way that many cultures have dealt with this problem is by providing *social assurance* – the collective enforcement of honest dealings that promises to punish dishonest conduct in greater magnitude than the reward such conduct would provide. Social assurance is the Hobbesian leviathan that brings all individuals' interests into alignment, at least in regard to certain crucial exchanges. In this way, even when trust is lacking, people can behave as if they trust each other, because of the overarching incentive structure provided by laws or other customs.

Social assurance. Any modern economy depends on a high level of certainty between participants that they will make exchanges honestly, but different cultures can differ in the relative importance of trust versus assurance. In a study of Japanese and American participants,

Japanese respondents reported lower levels of interpersonal trust, and assigned less importance to reputation and honesty than American respondents did, but there was no difference in the amount of caution the two populations believed was necessary in dealing with other people. The researchers concluded that this was due to the greater reliance on networks of committed relations in Japanese culture, which provides a strong incentive for people to behave in what amounts to a trustworthy manner in their exchanges with others (Yamagishi & Yamagishi, 1994).

There is also reason to believe that much of the prejudice and distrust that people have towards atheists (Edgell, Gereis, & Hartmann, 2006; Gervais, Shariff, & Norenzayan, 2011) is due to their relative lack of accountability to a higher power – their disbelief in a watchful and just God makes them targets of suspicion because they do not subscribe to the spiritual incentive structure that religious people do. Gervais and Norenzayan (2011) found that reminders of the effectiveness of secular authorities reduced distrust of atheists in religious people. This suggests that making effective sources of social assurance more salient can reduce the distrust people feel toward others. Similarly, countries with more effective secular systems of authority tend to have lower rates of distrust for atheists (Norenzayan & Gervais, 2015).

Distrust being an active state means it places demands on limited cognitive resources. In order to conserve these resources, people will be motivated to find reasons to trust others, or at least to limit their need to engage in distrusting cognitions. This implies that where one system of assurance is ineffective, people will seek another. For example, the degree to which one relies on God and governments for assurance appears to be hydraulic in nature – as people lose confidence in their government, they rely more on God (or the inherent goodness of others), and vice-versa (Kay, Shepherd, Blatz, et al., 2010). Assurance is a means by which we can produce an

environment where we can trust most people, most of the time. By having a system in place that governs everyone, we free ourselves from the need to be suspicious of every new person we interact with. It is only when we perceive some cue that social assurance is insufficient and that others may have some interest in exploiting or deceiving us that we activate distrusting cognitive processes.

Chapter 3

Cognitive Processing in a State of Distrust

Given that it is characterized as a response to a situation in which messages may be misleading, it should not be surprising that distrust seems to make people interpret information in a more critical, creative, and effortful manner. This has often been described as *non-routine* information processing (Schul et al., 2008; 2004).

Although this mode of cognition has some conceptual similarities with, and at times may intersect with central route processing as articulated in the Elaboration Likelihood Model (Cacioppo & Petty, 1984), it does differ in some important ways in terms of both mechanism and prediction. Central route processing is often understood as the cognitively expensive, effortful, rational evaluation of the presented message. Distrust is more defensive than it is truly rational – where central processing is characterized as evaluating the available evidence and the internal consistency of the message, (focused) distrust involves active resistance to persuasion, and is better described as a reluctance to accept the message and consists of seeking-out contradictions and an eagerness to accept ideas that oppose the one being presented.

While some experiments deal in focused distrust (Schul et al., 2004), many recent experiments concentrate on diffuse distrust – inducing distrust in one context, and then observing its effects in the context of a totally different task, free of contact with the person or stimulus that was used to induce the distrust in the participant (Schul, Mayo, & Burnstein, 2006; Mayer & Mussweiler, 2011; Posten & Mussweiler, 2013; Friesen & Sinclair, 2011). Several different methods have been used to provoke this state of distrust.

Inducing Distrust

Distrust researchers have used a variety of induction methods to manipulate trust and distrust in their experiments. Semantic primes are often used, though some methods of inducing distrust involve the participant in a situation that attempts to simulate the elements of a distrust-inducing encounter.

Semantic primes. Researchers often use semantic priming methods such as the scrambled sentence task (Bargh, Chen, & Burrows, 1996) or a modified lexical decision task (Dijksterhuis, Aaerts, Bargh, & van Knippenberg, 2000) to induce a state of distrust. The scrambled sentence method presents participants with a set of five words, arranged randomly, and requires them to use four of those words to create a grammatically-correct sentence. Researchers insert words relevant to trust or distrust into these sets in order to inconspicuously get participants to process these concepts, priming them toward a trusting or distrusting mindset (Friesen & Sinclair, 2011; Posten & Mussweiler, 2013; Mayer & Mussweiler, 2011).

Researchers have also used a modified lexical decision task, in which strings of characters are flashed on a computer screen very quickly, and participants must decide whether the characters form a word or not, to induce participants to be more trusting or distrusting. In order to induce these states of mind, researchers construct the program so that many or all of the real words are semantically close to “trust” or “distrust” (or their German equivalents, as is often the case; Mayer & Mussweiler, 2011; Posten & Mussweiler, 2013). As with the scrambled sentence task, this procedure works by activating the concepts of trust or distrust, which shapes the style of cognitive processing that follows.

Pattern detection/training inductions. Another strategy researchers have used to induce distrust is to put participants into a situation that simulates one of the main features of a real life

situation that would promote distrust, namely the need to determine whether or not another person is trying to deceive them. Posten and Mussweiler (2013) accomplish this by using an economic deception game, in which a computer was used to give participants the impression they were playing with another participant over a network. Participants in the *trust* condition were provided with advice from their "partner" that was in their best interest. Participants in the *distrust* condition were provided with advice that would result in them being shorted to the advantage of their "partner". Participants in the *control* condition were not given any advice throughout the task.

Another technique involves providing participants with collections of several question-and-answer exchanges to read, some of which have answers from actual women, while others have been answered by men trying to impersonate women. Participants in the distrust condition are asked to pick out the "impostors," and participants in the control condition are asked to rate how spontaneous each respondent was (Schul et al., 2004). Similarly, participants have also been trained to distrust specific faces and trust others through a task in which narrow-eyed faces are consistently paired with false statements, and wide-eyed faces are consistently paired with true statements (Schul et al., 2004). An abbreviated version of this 'eyes and lies' paradigm described previously in which participants were asked to memorize and form an impression of either a narrow-eyed (distrust-inducing) or round-eyed (trust-inducing) face, based on the idea that these are facial features that are generally perceived as negatively associated with trustworthiness (Zebrowitz, 1997).

The effects of these distrust manipulations have often been described with the blanket term "non-routine." Though the effects seem to be widely varied at a glance, they can be organized into a few related categories.

Non-routine Information Processing

From a very young age on through adulthood, we generally trust the things that others tell us to be true, particularly if we have no reason to doubt the information (Harris, 2012; Gilbert, 1991). In other words, to trust is 'routine'. However, when we find ourselves in a potentially misleading environment, we are more likely to take on non-routine ways of interpreting the information that comes our way. This is advantageous in a few different ways. First, we are more prepared for the unexpected. Second, we are more sensitive to non-routine actions, which could provide clues that we are being deceived. Third, when we behave in a non-routine manner ourselves, we become less predictable and therefore less controllable by others who may be trying to manipulate us through deception. Distrust has often been said to facilitate non-routine cognitions (Posten & Mussweiller, 2013; Schul et al., 2004; 2008), though this non-routine-ness apparently takes a variety of forms.

Stereotype reduction and difference bias. Posten and Mussweiller (2013) used an economic deception game to induce participants to either trust or distrust. After completing the induction task, participants were asked to provide feedback on a portrait of either a male or female person (disguised as a filler task). They found that distrust-primed participants judged female targets to be less stereotypically female than their counterparts in the control and trust-prime conditions, and even judged *male* targets to be *more stereotypically female* than controls or trust-primed participants, suggesting that male stereotypes were reduced under distrust as well. They found similar stereotype-reducing effects for ethnic and weight-based stereotypes. They found that this effect could be nullified if they had participants complete a similarity-detection task between the distrust prime and trait rating task, concluding that the observed

stereotype reduction was due to participants comparing the target to their stereotypes and seeing more difference than usual.

Simultaneous contradictory cognitions. Distrust has been shown to facilitate cognitions that conflict with the information provided, as if the participant is automatically preparing for the information to be false. Shul, et al. (2004) trained participants to trust wide-eyed faces and distrust narrow-eyed faces. Participants were then presented with the faces again, this time with a subliminal prime adjective followed by a target word that they were to identify as a noun or an adjective. On trials with a target *adjective*, it was either semantically congruent or incongruent with the prime adjective (target nouns were filler trials; virtually all nouns and adjectives are potentially congruent when paired together). Participants were faster to recognize the target adjective of an incongruent word pair (e.g., Hollow-Solid) when the words were presented with a distrusted face than when they were presented with a trusted face.

This activation of incongruent concepts has also been demonstrated when activating social categories in response to perceiving another person. Friesen and Sinclair (2011) primed participants with distrust or trust, and then showed them a video of either a White doctor or a Black doctor. The only group to show evidence of category activation were distrusting participants who watched the Black doctor video. They showed activation of both *Black* and *doctor* categories. This result is interesting, not only because it suggests that distrust can lead to the activation of multiple social categories pertaining to the same person, but also because the words used in the lexical decision task to represent each social category were often directly opposed to one another. Words in the doctor category included terms such as *health, intelligent, educated, caring,* and *compassionate,* whereas words in the Black category included terms such as *dealer, gang, crime, violent, aggressive,* and *dangerous*. So, in addition to finding that distrust

seems to lead to the activation of more categories, this finding also supports the notion that distrust facilitates the simultaneous activation of incongruent cognitions.

Schul and his colleages (2008, experiments 1 and 2) found that participants induced to be distrusting, either by means of a task asking them to discriminate gender imposters from genuine women, or to process an untrustworthy face, were faster to adapt to counter-intuitive prediction environments, in which higher scores on unnamed character traits predicted lower job performance.

***Creativity*.** Mayer and Mussweiller (2011) found that creativity could be enhanced by priming distrust. Creativity was judged in terms of the number of unique ideas, originality of ideas, and the number of different semantic categories used when freely generating alternative uses for household items. They found, decomposing their creativity measure into its component parts, that flexibility (defined as the number of semantic categories used) mediated the relationship between distrust and originality (judged creativity).

In sum, distrust has been found to increase a person's ability to detect differences and also their ability to think creatively and unconventionally, becoming more alert to non-obvious alternatives and possibilities.

Two Underlying Themes of the "Non-Routine"

The literature on the cognitive effects of distrust often makes the claim that a state of distrust facilitates non-routine cognitions, and indeed, the activation of adjectives of incongruent meaning to a prime word (Schul et al., 2004), reduced stereotypicality of trait ratings (Posten & Mussweiller, 2013), increased creativity (Mayer & Mussweiller, 2011), and enhanced performance on more complex puzzles (Schul et al., 2008) are all outcomes that differ from typical or routine patterns of thought. Then again, the phrase "non-routine" is vague and does not

provide information about what it is about distrust that produces these non-routine cognitions. The term is also unhelpful in that it does not imply specific outcomes. "Non-routine" can potentially describe a wide variety of outcomes, and leaves too much room for interpretation. After all, most psychology experiments involve many behaviors that would be considered non-routine by most people. It appears to me that laboratory studies involving distrust can be explained better by attributing distrust effects to two mechanisms – increased cognitive flexibility and contrast sensitivity.

Cognitive flexibility. When we are distrusting of our surroundings, it makes intuitive sense that we ought to be more open to a wider variety of interpretations of the information presented to us. This may help us to detect another person's attempts at deception, or in a more impersonal situation, allow us to regain our bearings after being disoriented. In other words, we become more flexible in our thinking – we rapidly activate multiple categories, connect the stimuli we perceive to more different categories and associates, and entertain more creative interpretations of the things we see and hear.

This flexibility is on display in a series of studies conducted by Friesen and Sinclair (2008), in which distrust-primed participants who viewed a video of a Black doctor were faster to identify words associated with both Black people and doctors than control participants or participants who viewed a White doctor. In fact, distrusting participants viewing the Black doctor were the only participants to respond quickly enough to category-relevant words to even be considered evidence of category activation. It is difficult to explain why the doctor category was not still activated among distrust-primed participants who viewed a White doctor, but these findings still do appear to support the notion that distrust promotes the activation of multiple categories in response to a stimulus.

Another clear case is a series of studies reported by Mayer and Mussweiller (2011), in which distrust-primed participants were found to generate more creative ideas for alternative uses of common household items. Further exploration of the phenomenon revealed that the relationship between distrust and originality was moderated by cognitive flexibility, operationalized as the number of semantic categories participants used in their responses.

A third case that suggests that distrust promotes cognitive flexibility is an experiment in which distrusting participants were given a set of puzzles consisting of arithmetic problems made out of matchsticks. Participants were to change incorrect equations into correct equations by moving one matchstick and thus changing two characters in the problem. The routine set of problems only required participants to move a matchstick to one number to another number, changing only the values of different numbers (e.g. IX=X-III is solved: IX-XI-II). The non-routine set, however, required participants to remove a matchstick from one of the numbers and move it to one of the arithmetic signs in order to make the equation correct (e.g. XI=VI-IV is solved: X=VI+IV). Distrusting participants had a higher success rate for the non-routine items, and trusting participants had a higher success rate for the routine items (Schul et al., 2008, experiment 3). Because the non-routine problems required participants to consider both the numbers and the arithmetic signs for alteration, it suggests that participants who had greater success at these problems were attending to both categories of symbols presented to them, allowing them to imagine a greater variety of possible solutions. Although the processing of the problem can certainly be considered to be outside the routine, it seems to exemplify cognitive flexibility rather well, because it requires participants to consider a wider variety of possibilities when approaching the problem.

Together, these findings suggest that it is helpful to understand distrust at least in part, as a state that promotes cognitive flexibility. This implies that we might predict that people who are already thinking in a distrusting way may find it easier to imagine ulterior motives, side effects, and alternative courses of action that a more trusting person might dismiss without a second thought. This prediction differs from what we would expect if distrust effects were assumed to be merely due to use of central processing because this enhanced receptivity to selected subsequent messages relies on the activation of or search for new ideas, and not necessarily the actual critical evaluation of the arguments being presented, or of the alternative positions being generated.

Contrast sensitivity. It seems that in addition to making cognitions more flexible, a state of distrust could also promote an emphasis toward detecting inconsistencies – between stimuli in the environment, between the words and actions of others, and between categories and their exemplars. The utility of this feature of distrust is that it allows distrusting persons to prepare themselves for the possibility that the information they receive is misleading (Schul et al., 2004).

Schul et al. (2004) used an "eyes and lies" paradigm to induce distrust, and found that participants were faster to identify adjectives that were incongruent with a subliminal prime word when the pair of words was accompanied by a narrow-eyed face, and were faster to identify adjectives that were congruent with the prime word when a wide-eyed face was on the screen. In a second experiment, they used the gender imposter paradigm. Using the same adjective-identification task, they found the same result as in the first experiment – distrusting participants were faster to identify incongruent target adjectives and control participants were faster to identify congruent target adjectives. In both cases, it seems that when people are in a distrusting state, they are in a state of readiness to detect contrasting information.

The same research team reprised their "gender imposter" paradigm (Schul et al., 2008), and found that participants in the distrust condition were faster to learn a counter-intuitive prediction pattern – in which higher scores on important job-relevant characteristics predicted *worse* job performance – than participants in the control condition. The same result was found with a second version of the prediction task, in which outcomes were determined by either the greater (routine) or lesser (non-routine) of two predictors. Distrusting participants were faster to catch on to the prediction pattern in which the lower of the two predictors influenced the outcome than participants in the control condition.

Because these studies were designed to demonstrate non-routine ways of processing information, these studies do not seem to directly support the idea that distrust necessarily results in contrast sensitivity. Nevertheless, when Schul et al.'s (2008) findings are understood in the context of other literature on the topic, it seems that this instance of non-routine information processing is really a finding of contrast sensitivity. In order to adapt to these abnormal prediction patterns faster than their trusting peers, distrusting participants would have to be more ready to perceive an abnormal pattern in which outcomes followed predictor variables in a different way than is usually expected. In other words, they were more ready to perceive a different kind of pattern.

Posten and Mussweiller (2013) found that a state of distrust can reduce stereotype-based judgments of individuals regarding gender, ethnicity, and weight. They then found that this effect could be eliminated by giving distrust-primed participants a task in which they listed the similarities between two sides of a picture, so it appears that the decrement in stereotype-based judgments is due to a focus on how the individual being judged differs from the stereotypes associated with their social group.

This use of a similarity-seeking task to nullify the contrast sensitivity enhancing effects of distrust serves as a particularly strong argument for contrast sensitivity as a basis of many of the effects of distrust previously described simply as "non-routine". The fact that the mechanism has been identified and can be interrupted holds much promise for future research into the mechanisms behind distrust's effect on information processing.

Interestingly, the study by Friesen and Sinclair (2011) referenced in the above section is relevant to both the cognitive flexibility and contrast sensitivity dimensions of distrust. Since they found that distrusting participants activated both *Black* and *doctor* categories after watching a video of a Black doctor, while all control participants and distrusting participants viewing a White doctor showed no evidence of category activation at all, this not only supports the idea that distrust promotes the activation of multiple categories, but due to the fact that the words used to represent the *doctor* category were also overwhelmingly positively-valenced, and the words used to represent the *Black* category were overwhelmingly negatively-valenced, their findings also imply that the distrusting participants were more sensitive to contrasting information.

Given the variety of ways that distrust-primed participants have found contrasts – detecting contrasting adjectives more quickly, contrasting an individual with their social category, activating both positively and negatively-valenced categories when presented with a stimulus – it seems that the contrast sensitivity prompted by distrust is rather general, and that distrusting people in real-world situations may find the contrasting information they are looking for just about anywhere. They may find contrary information in the message itself, in remembered anecdotes, or possibly even in their imagination.

This process can be distinguished from biased central route processing, at least as it is typically understood, in that it predicts reactions to subsequent messages. This feature of distrust is not incompatible with elaboration likelihood processes, but it does look beyond the predictions of that model temporally.

Next Steps

Breaking down the predicted nonroutine cognitive effects of distrust into these two main categories of effects – cognitive flexibility and contrast sensitivity – affords us the opportunity to make further predictions about how these changes in cognitive processing might translate to changes in how information is interpreted, and ultimately in how attitudes are developed and changed in response to that information. The extension of this theory from cognitive processes to one's beliefs about the world will be the focus of the experiments in this dissertation.

Chapter 4

Implications for Attitude Theories

The body of research concerning distrust and cognition up to this point has addressed socially-relevant topics such as stereotyping (Posten & Mussweiler, 2013; Friesen & Sinclair, 2011), and creativity in the workplace (Meyer & Mussweiler, 2011), but the effects of distrust on more complex information-processing tasks such as deciding whether or not to support a cause or seek a particular form of medical treatment over other options have yet to be investigated. Now that the research into the basic cognitive mechanisms of distrust has uncovered some of its basic characteristics, it seems that the time is ripe to apply these findings to activities that are closer to what people experience *in vivo*. How does enhanced contrast sensitivity change the way a citizen votes? How might increased cognitive flexibility change the way a person evaluates their health care options? In order to understand the social consequences of the cognitive features of distrust, we first need to understand the role they play in shaping attitudes and, ultimately, behaviors.

One informative example of the relationship between distrust and attitude change is the current public health issue of vaccine refusal. The rate of nonmedical exemptions from childhood immunizations has been on the rise in recent years (Omer, Richards, Ward, Michelle, et al., 2012). Lack of trust in physicians and the healthcare system in general has been indicated as an important contributor to this increase (Benin, Wisler-Scher, Shapiro, et al., 2005). Anti-vaccination websites tend to promote this distrust, emphasizing the notion that medical professionals are ignorant of the facts, incompetent, or actively deceiving their patients, while actively promoting naturopathy, homeopathy, and other complementary/alternative medicine (CAM) treatments and forms of prevention (Wolfe, Sharp, & Lipsky, 2002; Zimmerman, Wolfe,

Fox, et al., 2005), the use of which has been linked both to vaccine refusal and to the acquisition of vaccine-preventable illnesses (Downey, Tyree, Huebner, et al., 2010; Salmon, Haber, Gangarosa, et al., 1999).

Like much of the existing distrust research (Posten & Mussweiler, 2013; Meyer & Mussweiler, 2011; Schul et al., 2008), the present investigation conceptualizes distrust as a diffuse state, and not necessarily as a specific attitude toward a specific person or entity. Distrust is essentially the feeling that the environment may not behave quite in the way you would normally expect – a person may be telling you falsehoods; a company might have undisclosed, sinister motives for working with you; a familiar walking surface may suddenly be slippery. In each of these situations, you have a feeling that each action you take requires more thought than usual, and is accompanied by some degree of uncertainty about the consequences that will come of it. And in each case, this feeling and this way of processing information would still be operating even if you were accompanied by your most trusted friend.

Because the concepts of trust and distrust are directly related to one's degree of certainty about objective reality, the effects of distrust may be less likely to apply to messages regarding matters of opinion or of cultural values and norms, such as support for same-sex marriage or abortion. Distrust's effects should be most robust when it comes to questions for which there can be such a thing as an expert consensus, such as the efficacy and safety of vaccines, the shape of the earth, or the evolution of humans.

Furthermore, based on these cognitive influences on attitudinal processes, distrust should usually bias people against that expert consensus. Because the consensus position will shape nearly any conversation on the topic, it will be the most available position against which to search for contrasting information, and to which alternatives will be sought. One rarely needs to

justify adhering to mainstream medical advice – following the normal vaccination schedule, for example, is the default. But a state of distrust carries with it the implication that assumptions should be challenged, making it harder to accept arguments from authority, even when that authority is legitimate.

Conceptualizing distrust as a state in which a person activates a wider variety of categories in response to a stimulus, as evidenced by studies demonstrating that distrust resulted in the activation of more social categories (Friesen & Sinclair, 2011) and more unique uses for common items (Mayer & Mussweiler, 2011), and in which a person will be more sensitive to dissimilarities within their environment (Posten & Mussweiler, 2013; Schul et al., 2004), has some potentially useful applications to attitude change and persuasion theories. Whereas a distrusting person is colloquially believed to be relatively resistant to persuasion and more cautious in general, breaking distrust down into specific mechanisms may help us to understand more specifically why this pattern of cognition appears to be the norm, and in what ways distrust is unique among routes to resisting attitude change. Given what is already known about how distrust impacts cognition, it seems likely to change the way people respond to both the source of an argument, and the argument itself.

Chapter 5

Distrust and the Source of a Persuasive Argument

The source of a persuasive message plays an important role in determining its effectiveness. We are not likely to be persuaded by people who appear to lack expertise or who we believe to be biased by their own personal motives (Eagly, Chaiken, & Wood, 1981). Person-specific distrust is an obvious barrier to being persuaded by a particular person, but the cognitive patterns associated with a diffuse state of distrust (Schul & Peri, 2015) may also influence how persuasive we find *anyone*, even if we have no prior reason to distrust them personally. The increased cognitive flexibility and contrast sensitivity that have been observed in tests of basic cognitive functioning may extend to how we develop our beliefs by making other people appear more biased, by devaluing expertise, and by increasing the credibility of fringe figures.

Cognitive Flexibility and Attribution

Attributional approaches to persuasion (Eagly, Chaiken, & Wood, 1981) indicate that people tend to have expectations about what a given communicator might say before he or she says it, and that the more the communicator's message conforms to these premessage expectations, the more that communicator will be considered to be biased, and thus less credible (unless the position can be accounted for by the communicator's expertise). Those communicators who contradict a person's premessage expectations tend to gain extra credibility with that person. The fact that distrust appears to facilitate the activation of more and more varying social categories (Friesen & Sinclair, 2011) seems to indicate that a distrusting person will be more likely to believe that a communicator is biased (because they have more expectations available in memory). The activation of more social categories should result in a

wider variety of premessage expectations. Therefore they will be less convinced by what a credible communicator says than he or she would be otherwise.

For example, given the findings of the distrust literature that a state of distrust promotes the activation of multiple simultaneous categories in response to a single stimulus (Friesen & Sinclair, 2011; Mayer & Mussweiler, 2011), it is possible that a distrusting message recipient will activate a wider variety of categories in response to reading a message from a medical doctor, some of which may even be contradictory. Even categories with conflicting stereotypes may be activated simultaneously – the doctor may also be seen as a business owner, an associate of large and powerful pharmaceutical companies, or even an agent of a religious group or the government. This wide range of activated categories would be liable to lead to a considerably wider range of premessage expectations, and more reasons why the doctor may be motivated to distort or withhold some of the facts in the eyes of the message recipient.

This notion that one's own distrust would make others seem to possess a broader range of personal biases corresponds with a basic element of distrust – the suspicion that the other person has ulterior motives, or goals that are incompatible with your own that may poison the exchange.

It should certainly be no surprise that a person who is in a state of distrust would find it difficult to take another person's words at face value, since that is what anyone would think of when they hear the word "distrust," but this cognitive flexibility hypothesis may help clarify the mechanism behind that disinclination to change one's position that is so characteristic of the concept of distrust - the pre-activation of more social categories makes it easier to dismiss other people's opinions as biased.

Contrast Sensitivity – Reduced Impact of Good and Bad Qualifications

A curious side-effect of distrust's facilitation of difference-finding is that it appears to actually reduce stereotyped judgments. Posten and Mussweiler's (2013) research into the topic has demonstrated this by showing that participants primed with distrust rated Turkish, female, and overweight individuals lower on their respective stereotypical traits. Presumably, this is because when a person is being judged according to a particular trait, they are being judged *relative* to that trait. A dissimilarity-oriented approach thus makes it seem as if the target and the standard are further apart.

What this seems to say for communicators is that characteristics that increase or decrease their credibility to a typical audience will have a *weaker* impact when addressing a *distrusting* audience. Experts will be judged to be less expert-like. "Quacks" (people who do not have the legitimacy of a doctor, but espouse/have the trappings of being a doctor) will be judged less quack-like.

However, this prediction depends a great deal on context. If an expert and a quack are presented together, and the viewer is more prepared to find dissimilarities, then it follows that the two would be judged to be *more* different from each other, and the discrepancies in their levels of expertise would become *more* apparent.

So, distrust should make the difference between experts and sham artists appear smaller than they normally would, *when they are each presented in isolation*, because the viewer will look for differences between the individual and their most salient social group. That same difference between the two individuals will seem larger than it normally would *when they are presented together*, because the viewer will have the opportunity to look for differences between the two individuals.

Still, at least in the case of the vaccination debate, proponents and opponents rarely do occupy the same space. It seems that distrusting people who aren't sure about whether to vaccinate will have an increased vulnerability to quackery and an increased resistance to qualified experts at the same time, elevating the likelihood of vaccine refusal.

If this prediction bears out, it could have interesting implications for public debates about vaccination and other science-based policies. Many experts avoid having face-to-face debates with quacks because they believe it will only lend legitimacy to the quack (To paraphrase Richard Dawkins, it would look great on the quack's CV, but not so good on his own). This may hold true for most of the viewing public, but for the subset who are perhaps more prone to distrust, having the opportunity to make side-by-side comparisons may allow their contrast sensitivity to serve an adaptive purpose, and make it more clear which party is qualified to speak on the issue.

Chapter 6

Distrust and the Issues or Arguments at Hand

Whether or not we are suspicious of the person who communicates a message to us, we may be suspicious of the information itself. It may be second-hand information from an untrusted source, it may share features with other messages we have found to be suspicious, or we may have been recently primed to take a distrusting approach to stimuli. In any of these cases, distrust would lead us to interpret messages in different ways than we would if we were feeling more trusting. Based on what we know about the cognitive characteristics of distrust, it should lead people to entertain more alternatives to what they are being told, and to put more effort into searching for contradictions.

Cognitive Flexibility – What Else Might Be True?

Increased cognitive flexibility that comes with a state of distrust is bound to lead us to entertain different ways of interpreting the content of an argument, possibly even going beyond the argument itself in an effort to devise what relevant information might not be present in the argument. In particular, a distrusting person will be interested in 1) what other outcomes are likely to result from the recommended course of action (i.e. “side effects”), and 2) what other actions are available to him or her that could also result in the desired outcome (alternative actions).

This pattern can be observed among people suspicious of vaccination. The fear of side effects such as autism, autoimmune disorders, or even diabetes pervade discussions about the safety of vaccination, and antivaccination advocates often promote alternative, often ineffective, methods of avoiding viral diseases such as hand washing, homeopathic treatments, and proper nutrition (Wolfe et al., 2002).

Contrast Sensitivity – A Reason To Disbelieve.

We can also expect people in a distrusting state to be more vigilant in their mental search for contradictions with the position being presented to them. Distrusting people will more readily generate examples from as wider variety of sources in response to an argument than trusting people will. This critical examination is often helpful for protecting people in situations where misinformation abounds (Lewandowski, Ecker, Seifert, Schwarz, & Cook, 2012), but the same processes may promote an unhelpful level of resistance where the general consensus is legitimate.

This tendency is also on display in people who are suspicious of vaccines. Arguments against the safety and efficacy of vaccines are often supported by anecdotes of people getting sick despite vaccination, or suffering symptoms of a neurological disorder soon after administration of a vaccine (Zimmerman et al., 2005).

The Present Research

The present studies represent an effort to examine how contrast sensitivity and cognitive flexibility may change the way people respond to the content and source of persuasive messages. The first three studies will test my predictions about the contrast-sensitizing effects of distrust. Study 1 will address how this sensitivity to contrast may make statements that contradict previously presented positions seem more plausible by increasing the availability of opposing cognitions. Study 2 will address how sensitivity to contrast can make members of different professions seem more or less stereotypical of their group, and therefore more or less authoritative, depending on whether they are presented alone or with members of contrasting groups. Study 3 will then apply this principle directly to the elaboration likelihood model of

persuasion and examine whether the contrast-sensitizing effects of distrust lead to the generation of more counterarguments in response to an argument.

The final three studies will address the attitudinal implications of how cognitive flexibility and creative thinking are promoted by a state of distrust. Study 4 will address how distrust increases the accessibility of semantic categories, which may make it easier for people to generate opposing arguments to a position. Study 5 will demonstrate how distrust may make it easier for people to assimilate disparate concepts by adding intermediary cognitions, which is a likely engine for belief in, and maintenance of, conspiracy beliefs. Study 6 will test the prediction that distrust does not necessarily lead to a tendency to reject claims, and that arguments that take advantage of the contrast sensitization and cognitive flexibility that come with a state of distrust will be more effective when presented to distrusting persons.

This line of research will serve to apply the findings of previous research on the cognitive effects of distrust to the realm of persuasion, a necessary step toward using these findings for the public good, and will also help to advance and expand our understanding of the cognitive features of distrust.

Chapter 7

Pilot Studies

For a variety of reasons, it was necessary to perform a series of pilot tests on manipulations of trust and distrust before beginning with theory-testing experiments. I wanted to test four different methods of inducing a state of distrust, some of which involve semantic priming – word searches containing distrust- or trust- related words, and a classic scrambled sentence paradigm using the same distrust- and trust-related words. Others involve more realistic situations designed to promote distrust – an economic deception game (Posten & Mussweiler, 2013) in which a simulated partner behaves in either a trustworthy or untrustworthy manner, and an imposter-detection task in which participants in the distrust condition were asked to evaluate whether or not an interview subject was a genuine student athlete, or a non-athlete impersonating one.

Past studies have done little to demonstrate the validity of these methods. Direct measures of participants' mindset regarding trust or distrust are generally not included in these studies (Friesen & Sinclair, 2011; Mayer & Mussweiler, 2011; Posten & Mussweiler, 2013; Schul, Mayo, & Burnstein, 2008). Schul, Mayo, and Burnstein (2004) did validate their 'eyes and lies' induction technique – in which wide-eyed faces were consistently paired with true information and narrow-eyed faces were paired with false information – in a manipulation check where participants rated the trustworthiness of the different faces after completing the induction task. However, this measure does not fully validate the induction procedure, because the narrow-eyed faces were designed to *look* untrustworthy in the first place. Participants likely would judge these faces to be less trustworthy even without paired false information (Akehurst et al., 1996; Schul et al., 2008).

Many of the core studies that form the empirical basis for this dissertation were conducted in Germany (Posten & Mussweiler, 2013; Mayer & Mussweiler, 2011) or Israel, (Schul et al., 2004; Schul et al., 2008) with materials presented in languages other than English. The lone English language study was conducted in Canada (Friesen & Sinclair, 2011). So it was necessary to adapt prime words from their original languages. Cultural differences needed to be taken into account as well. A manipulation in which participants are asked to identify "gender impostors" posing as women (Schul et al., 2004; 2008), for instance, could be considered disrespectful when used in a participant population that is majority female and contains transgender people. To avoid this, the task was adapted to have participants identify impostor student athletes.

It also seems essential to validate priming methods, in particular, before using them in an experiment. Researchers frequently have been unable to replicate semantic priming effects (Bohannon, 2014; Cesario, 2014; Pashler, Rohrer, & Harris, 2013), calling into question the validity of these manipulations. On the other hand, recent meta-analysis has found that priming can be effective, at least when the target measure is both behavioral and directly related to the prime. Weingarten, Chen, McAdams, and colleagues (2016) conducted a meta-analysis specifically investigating the efficacy of priming studies in which the target of the priming manipulation were behavioral (not attitudinal or judgmental) in nature, and were directly (not abstractly) related to the target behavior (e.g. prime: words related to giving; target: amount of money given to another person). In these favorable conditions, they found that semantic priming can have an effect, especially if the prime is designed to activate a goal. It did not directly speak to priming studies that were non-behavioral or in which the target behavior was indirectly related to the prime. The distrust priming reviewed here and present in the pilot studies is directed

toward targets that are often indirectly related to the prime or else attitudinal rather than behavioral, so this series of pilots will also contribute to a broader perspective on this current controversy.

Finally, social cognitive experiments on the effects of distrust have yet to use any manipulation checks validating that distrust-related states were aroused. Thus, the work to date leaves open the question of whether distrust really is the construct being induced in those studies, or even what specifically is the state of "distrust" that causes observed effects in experiments. So, in addition to seeking replication of some consequences of inductions observed by other researchers, such as differences in similarity ratings of different objects (Posten & Mussweiler, 2013) and the ability to solve routine and non-routine matchstick puzzles (Schul et al., 2008), I used a variety of self-report distrust measures as well.

Outcome Measures from Other Studies

I tried to reproduce the effects of distrust manipulations on participants' judgments of the similarity of different objects observed by (Posten & Mussweiler, 2013) and their effect on participants' performance on object-manipulation puzzles observed by (Schul et al., 2008). These findings have contributed to establishing the notion that distrust leads to non-routine processing of information, but were performed in non-English-speaking populations in different countries. It was therefore important to determine whether the same effects would be observed in American participants.

Similarity Judgments. Posten and Mussweiler (2013) have found that participants judged pairs of objects to be less similar after their distrust induction, implying that a state of distrust facilitates the perception of contrast. Each of the following distrust induction methods

were evaluated using an adapted version of this measure, using English rather than German word pairs.

Matchstick puzzles. Schul, Mayo, and Burnstein (2008) found that distrust-primed participants were able to solve a greater number of matchstick puzzles that required more complex thinking to solve, though they had no such advantage with the simpler puzzles. They interpreted this to mean that distrust promotes non-routine cognition, in which people think more unconventionally in response to feeling subject to the possibility of being fooled. This measure was reproduced in these pilot studies in order to see if our adaptations of their distrust manipulations would have the same effect.

General Trust and Distrust

In order to directly measure participants' subjective feeling of how trusting they felt toward others, and how generally distrusting they were as people, direct self-report measures were used.

Interpersonal Trust and Caution. General trust for others and interpersonal caution was measured according to Yamagishi and Yamagishi's (1994) two-factor scale. The *trust* factor of the scale assesses confidence in the beneficence of other people's intentions, and the *caution* factor assesses the felt need to be wary of others' motives.

Mistrust. The IPIP mistrust scale (Goldberg et al., 2006) was also used to assess general distrust, or suspicion of other people. In order to make this measure more reflective of present, in-the-moment distrust, participants were instructed specifically to respond to the questions as they felt in the present moment.

Domain-Specific Trust/Distrust

Trust for specific others was also measured in order to see if the various trust manipulations would change participants' trust for particular kinds of people. Participants were asked to report how much they trusted their neighbors, relatives, self, science, police, and religion.

Uncertainty and Lack of Control

The concept of distrust is closely related to feelings of uncertainty and lack of personal control. Research on control deprivation has revealed similar effects to those that have recently been associated with distrust – increased information-seeking, increased effort in processing information, and more conservative, cautious evaluation of information (Pittman & D'Agostino, 1985, 1989; Pittman & Pittman, 1980).

Variety-Seeking and Intellectual Openness

Based on the findings that distrusting participants were faster to identify words that conflicted with prime words (Schul et al., 2004), came up with more creative uses for common items (Mayer & Mussweiler, 2011), and judged members of social groups in less stereotypical terms (Posten & Mussweiler, 2013), it was hypothesized that distrust may be associated with greater open-mindedness and a preference for novelty. The variety-seeking and intellectual openness scales found on the international personality item pool (IPIP; Goldberg et al., 2006) were used to try to capture these effects. As with the IPIP mistrust scale, participants were instructed to answer the questions based on how they presently felt, not necessarily as they felt generally.

Pilot 1A Word Search Pilot – Student Sample

Semantic primes have been one of the main ways that experimenters have promoted a distrusting state of mind in participants (Friesen & Sinclair, 2011; Mayer & Mussweiler, 2011; Posten & Mussweiler, 2013). A word search puzzle has not yet been specifically used to promote distrust, but it is an established vehicle for semantic priming (Bargh, Gollwitzer, Lee-Chai, Barndollar, & Trötschel, 2001). Its reliability, however, has been questioned as of late due to some documented failures to replicate (Harris, Coburn, Rohrer, Pashler, 2013). These concerns make an exploratory study essential before conducting experiments using the manipulations. Two studies were conducted using an interactive word search as a method of delivering semantic primes for trust and distrust. One used a university student sample, the other a sample of MTurk workers – online laborers in Amazon's crowdsourcing marketplace platform.

Participants who received the distrust-prime words were expected to score higher on measures of distrust, variety-seeking, openness, caution, and uncertainty than participants who placed in the no-prime (control) condition. They were expected to score lower on measures of trust, and were expected to solve more of the matchstick puzzles that required more complex manipulation of the matchsticks (nonroutine cognitions), and solve fewer of the puzzles that required more simple manipulation of the matchsticks (routine cognitions).

Method

Participants and Design

Of all 201 participants, 30.3% were male, 68.7% were female, and 0.5% were transgender. They were 60.7% White, 8.5% Black, 12% Asian, and 17% Latino. The 3.5% who selected "other" in response to the question identified themselves as Indian, Indo-Guyanese American, Sri Lankan, European American, and American. Ninety percent of participants

identified as either lower-middle class, middle class, or upper-middle class. Economically, .5% identified themselves as impoverished, 3.5% lower class, 17.2% lower-middle class, 44.9% middle class, 29.3% upper-middle class, 3% upper class, and 1.5% super rich.

This pilot study was a between-subjects, two-condition experiment. Participants were randomly assigned to the *distrust* or *control* condition.

Materials

The two levels of the independent variable were carried out by providing links from within the Qualtrics survey to interactive word search puzzles created by the researchers and hosted on the website Teachers-direct.co.uk. Each word search contained the neutral words *entertain, plaster, varnish, grapes, stride, llama, ocean,* and *tea,* in addition to eight condition-specific prime words. In the *distrust* condition, prime words were *distrust, fake, doubt, suspicion, uncertain, questionable, suspicious,* and *mistaken*. *Distrust* prime words were adapted from the prime words used by Posten and Mussweiler (2013) in their scrambled sentence task. The *control* condition contained eight additional words that were not intended to prime any particular concept: *apartment, training, elephant, oldest, excited, bookish, swift,* and *fancy*. Every effort was made to create the easiest, least frustrating puzzles possible. All letters were set to uppercase, words were spelled only in natural (left-to-right or top-to-bottom) directions, and the un-used letters were put on the "easy" setting provided by the website. The link to the website opened in a new window so that it would be easy for participants to stay on the same survey page while completing the puzzle.

Measures

Confidence and Personal Control. Two single-item questions (Appendix A) were adapted from research on depression, uncertainty, and personal control. The items, "at this

moment, I feel confident in the correctness of my opinions, judgments, and decisions," and "at this moment, I feel I have sufficient personal control and power over the things that happen in my life" were taken verbatim from Weary and colleagues' (1993) work on depression, control motivation and social information processing, and were used to measure confidence and personal control. These items have generally been assumed to capture fairly stable states of certainty and judgmental confidence, and the extent to which a person feels events are under their control or not.

IPIP Mistrust. The six-item "Mistrust" questionnaire from the International Personality Item Pool (IPIP; Goldberg et al., 2006; Appendix B) was used to measure general mistrust for other people. Participants were instructed to "Describe yourself as you generally are now, not as you wish to be in the future," in order to encourage participants to report their current state of mistrust rather than their more stable, dispositional levels of mistrust, which is the way the items are typically used. The instrument is six items long, two of which are reverse-scored before items are averaged. Each item was measured on a five-point Likert-style scale (1 = "Very Inaccurate – 5 = "Very Accurate). In community samples, inter-rater reliability was high (α=.83; Goldberg et al., 2006). Higher scores indicate *more* mistrust.

Trust For Specific Other Persons and Institutions. In a set of questions adapted from Goertzel's (1994) research on the relationship between trust and belief in conspiracy theories, participants were asked to report how much they trusted their neighbors, relatives, self, science, and religion (Appendix C). Goertzel's original investigation only asked about trust in the police, their neighbors, and their relatives. He found that these measures had relatively low inter-item reliability (α=.57), but correlated significantly and positively with belief in conspiracy theories. Responses were recorded on a 5-point scale (1 – strongly disagree to 5 – strongly agree)

Trust and Caution Toward Others. A six-item scale measuring participants' perception of the trustworthiness of other people was included, as was a seven-item scale to measure the amount of caution participants felt was necessary to use around other people (Yamagishi & Yamagishi, 1994; Appendix D). Responses were recorded on a 5-point scale from "strongly disagree" to "strongly agree." Reliability has been found to be moderate for both factors (α ranged from .70 to 7.8 for general trust, and .61 to .74 for caution; Yamagishi & Yamagishi, 1994). These two measures are generally used to measure dispositional distrust and have been found to correlate positively with the use of negative hypothesis testing in evaluating initial hunches (Mayo, Alfasi, & Schwarz, 2014), but participants in these pilots were instructed to "Describe yourself as you feel at this very moment, not necessarily as you feel generally."

Matchstick Puzzles. A series of mathematical puzzles illustrated in the form of matchsticks was created to replicate Schul, Mayo, and Burnstein's (2008) test of nonroutine cognition after distrust inductions. In all, there were eight mathematical equations written in Roman numerals, as opposed to the Arabic numerals that participants are presumed to be more accustomed to. The numerals were shown to participants in the form of arrangements of matchsticks. Each equation was incorrect. For each equation, participants were asked to make the equation correct by mentally moving one matchstick from one symbol to another, thus changing both symbols (Appendix E). They were to type the answer they came up with in a text field.

Of these eight equations, four of them could be solved by moving a matchstick from one numeral to another, and the other four could only be solved by moving a matchstick from a numeral to one of the operational symbols (+,-,=). Performance was scored 0-4 for each set of puzzles, one point for each correct answer. Participants were given three minutes to complete the task.

Similarity Judgments. Fifteen pairs of nouns were presented to participants, such as *Lake and Ocean, Lion and House Cat,* etc., with the request that they judge how similar or dissimilar the two things were on a six-point scale (Posten & Mussweiler, 2013; Appendix F). These ratings were averaged for each participant, with higher scores indicating a greater tendency toward judging objects to be more similar.

IPIP Variety and Openness. IPIP scales for variety-seeking (Appendix G) and intellectual openness (Appendix H) were used (Goldberg et al., 2006). Both of these instruments are measured using a five-point Likert-type scale from "very inaccurate" to "very accurate." The variety-seeking scale consisted of 10 items, three of which are reverse-scored. Inter-item reliability was adequate in a community sample (α=.80). The intellectual openness scale was also 10 items long, with five reverse-scored. It also demonstrated adequate inter-item reliability in the same community sample (α=.80).

Demographics. Participants were asked their gender, race/ethnicity, and the economic class they identify with (Appendix I).

Stance on Childhood Vaccinations. Participants were also asked whether they believed children should receive *all* recommended vaccinations, only *some* of the recommended vaccinations, or *no* vaccinations at all (Appendix J). This measure was included for use in a different study.

Procedure

Participants were recruited online from the University of Connecticut Psychology participant pools of either the Storrs campus, or the pool covering all of the regional campuses. Storrs participants reserved a time and received a link to the study via email at the designated time. Regional campus participants followed the link on the pool website to get to the study.

The study was administered electronically and distributed online. After reading the information sheet, the program randomly assigned the participant to the *distrust* or *control* condition, providing a link to the appropriate word search. The interactive word search puzzles displayed a particular animation upon completion. In order to confirm that participants had completed the word search, they were asked on the following page of the survey to briefly describe what they had seen when they completed the puzzle. After this manipulation was complete, participants were introduced to the measures described above in that order.

Throughout the survey, items serving as *attention checks* were included. These included an item included on the same page as the questions concerning trust for specific others that read, "Please answer 'Disagree' for this item," an item embedded in the page of similarity ratings that asked participants to judge how similar water was to water (the correct answer was "Very Similar"), and an item included among the IPIP Variety questions asking participants to answer, "neither accurate nor inaccurate." These questions were included with the goal of identifying participants who were going through the experiment without reading the items.

At the end of the survey, participants were presented with the debriefing statement, which revealed to them the full purpose of the experiment, the intent behind the priming task, and the hypotheses of the study. After the debriefing statement participants were given the opportunity to provide open-ended feedback to the question, "Now that you have finished with this study, do you have any comments for the researchers? Did anything seem off? Do you have any ideas for improving the study's implementation?" Participants were thanked for their participation.

Results

Data Preparation

Total time spent on participation was calculated by subtracting the starting time from the ending time. Test runs through the experiment took approximately 11 minutes to complete, so participants were excluded from analyses if their total time was less than nine minutes, or more than thirty minutes (included participants took an average of 21 minutes to complete the study, SD = 5.19). Participants were also excluded if they failed one or more of the attention checks, or if they did not indicate any knowledge of the completion animation of the word search, which was an explosion of letters followed by a green check mark and the words, "puzzle complete." Responses were judged to indicate non-completion if they gave no answer, reported skipping the puzzle, or otherwise failed to mention any features of the completion screen. Of the 201 total participants, 107 survived the screening process. Because this was a pilot study, these stringent criteria were used in order to make sure, to the highest degree possible, that the participants included in analysis had actually gone through the manipulation, and that the priming effect of the word search itself was being evaluated.

Analyses

IPIP Mistrust. Distrust-primed participants scored lower on the IPIP mistrust scale (M = 2.37, SD = .646) than control participants (M = 2.62, SD = .728) $t(199) = -2.564$, $p = .011$, $d = .12$, suggesting that the distrust prime actually *reduced* feelings of mistrust.

Trust For Specific Other Persons and Institutions. Similarly, distrust-primed participants reported trusting their relatives *more* (M = 4.31, SD = .776) than control participants did (M = 3.95, SD = 1.041), $t(105) = -2.006$, $p = .047$, $d = .39$.

There were no other significant effects of the distrust manipulation. Descriptive statistics and outcomes of all t-tests are reported in Table 1.

Based on these results, it does not appear that the word search priming method will be a good candidate for studying the effects of distrust. The manipulation had no effect on the majority of the measures, and if anything, made participants report a more *trusting* attitude instead of a distrusting one. Scores were lower on the IPIP measure of mistrust, and higher for trust in one's relatives. It's possible that greater general distrust drives someone to trust their own family more, but the lower score for general mistrust conflicts with such a conclusion.

Discussion

These findings seem to suggest that this method of priming for *distrust* may only activate *trust*. Because the processing of negatives requires more work than processing the root concept itself, and semantic priming works by means of implicit processes, it could be that the root concept (trust) is activated without being negated (see Gilbert, 1991). Further studies that include a *trust* prime condition may help to clarify this issue.

Pilot 1B Word Search Pilot – Mturk Sample

Method

Participants and Design

Sixty participants were recruited via Amazon.com's Mechanical Turk (MTurk) – a service that offers small payments for people who complete short tasks over the Internet. Of the 60 participants in the MTurk sample, 83.3% were White, 10% Black, 6.7% Asian, and 3.3% Latino. Economically, 5.6% identified as impoverished, 22.2% lower class, 27.8% lower-middle class, 33.3% middle class, and 9.3% upper-middle class.

This pilot study was a three-condition experiment. Participants were randomly placed in the *trust, distrust* or *control* condition. The word search for the *trust* condition contained the prime words *trusting, credible, authentic, dependence, certain, reliable, honest,* and *sympathetic*. The *distrust* and *control* primes were the same as in the student sample version of this pilot.

Participants were included in analyses if they met the criteria described in the student pilot described previously, and 54 met full inclusion criteria. On average, included participants completed the experiment in 18.46 minutes (SD = 4.72).

Materials

The same study materials were used for this sample as for the student sample.

Measures

The measures were identical to those used in the UCONN student sample, with the exception that the gender item was unintentionally omitted from the survey questions.

Results

Analyses

Fifty-four participants met full inclusion criteria. One-way between-subjects analyses of variance (ANOVAs) were calculated to examine differences in each of the dependent variables between the *trust, distrust,* and *control* conditions. No significant effects of condition were observed for any of the dependent measures. Solved non-routine matchstick puzzles did show a marginal effect $F(2, 51) = 2.24$, $p = .081$, in which *control* ($M = .53$, $SD = .96$) participants solved slightly more of the four puzzles than did *trust* primed ($M = 0.00$, $SD = 0.00$) and *distrust* primed ($M = .28$, $SD = .83$) participants. Effect sizes tended to be small on most measures ($\eta^2 = .02$ or lower), with the exception of performance on the non-routine matchstick puzzles, which had a medium effect size ($\eta^2 = .08$). Tests for all variables for this study are reported in Table 2.

Post hoc exploratory analyses contrasting the combined prime conditions to the control condition also failed to show a significant effect of priming.

Discussion

This pilot did not provide any evidence that the word search priming procedure effectively influenced MTurk workers' state of trust or distrust. The unanimously non-significant comparisons also yielded overwhelmingly small effect sizes, with the exception of the number of solved non-routine matchstick puzzles.

Pilot 2A Scrambled Sentence Task Pilot – Student Sample

The scrambled sentence task has been used in the past to prime distrust specifically, but only in Germany, using stimuli presented in German (Mayer & Mussweiler, 2011; Posten & Mussweiler, 2013). Priming as a general procedure has of course been used with English words, conducted with U.S. samples (Bargh et al., 1996), but has rarely been used to prime distrust specifically. So again, pilot studies using both student and MTurk populations seemed essential.

Method

Participants and Design

All 122 participants were recruited online from the University of Connecticut Psychology participant pool. Of this sample, 68% were White, 3.3% were Black, 22.1% were Asian, and 10.7% were Latino. The 3.3% who provided text responses to the question identified themselves as Middle Eastern, Native American, and Nepali. Economically, 3.3% identified themselves as lower class, 16.5% lower-middle class, 38.8% middle class, 34.7% upper-middle class, and 6.6% upper class.

This pilot study was a three-condition experiment. Participants were randomly placed in the *trust, distrust* or *control* condition.

Inclusion in analyses was determined the same way as it was in the word search studies, with the exception that participants were not given the completion test designed for the word search task. For this study, 114 participants met full inclusion criteria. On average, included participants finished the experiment in 16.91 minutes (SD = 4.79).

Materials and Procedure

Fifteen scrambled sentences with prime words embedded in them were presented at the beginning of the Qualtrics survey. This manipulation was adapted from that used by Posten and Mussweiler's (2013) English translations of the German stimuli they used in their study. Some of the German words had two comparable English word translations. These words were replaced with the nearest one-word English alternative. Prime words for the *trust* condition were *trusting, credible, authentic, dependence, certain, reliable, honest,* and *sympathetic*. For the *distrust* condition, they were *distrust, fake, doubt, suspicion, uncertain, questionable, suspicious,* and *mistaken*. For the *control* condition, there were no prime words, but the words *apartment, training, elephant, oldest, excited, bookish, swift,* and *fancy* were put in the place where the prime word would have been (Appendix K).

The study was administered electronically over the Internet. Directly after reading the information sheet, the program randomly assigned participants to the *trust*, *distrust*, or *control* condition and presented them with the corresponding version of the scrambled sentence task. In this task, participants were presented with 15 sets of five words, randomly arranged. They were instructed to mentally remove one of the words and make a grammatically correct sentence out of the remaining four words (the string *stay doubt to they had*, for example, could be solved by either "they had to doubt" or "they had to stay."). After this task was complete, the rest of the procedure was identical to that of the previous word search studies.

Results

The only standard for inclusion in analyses that yielded even one significant result was the most stringent – participants who completed the manipulation, passed all attention checks, and finished the study within two standard deviations of the average completion time (longer than 4.66 minutes, and shorter than 31.32 minutes) were included. Between-subjects ANOVAs were performed for each dependent measure to detect differences between the three conditions.

In this sample, there was a significant effect of condition on reported trust in relatives $F(2, 110) = 3.35$, $p = .039$, $\eta^2 = .06$. Post-hoc analysis using the Bonferroni adjustment revealed that trust-primed participants reported trusting their relatives significantly more ($M = 4.62$, $SD = .54$) than did control participants ($M = 4.26$, $SD = .72$), $p = .034$, $d = .57$. Tests for all measures are reported in Table 3.

Discussion

The fact that the distrust prime did not significantly affect this measure while the trust prime did, may suggest that my hunch is correct that the inherent negation of the concept of "trust" in priming distrust poses problems from a cognitive processing standpoint. However, the fact that this is one significant finding among a large number of measures, and could only be found using stringent inclusion criteria should make one cautious in drawing any broad conclusions.

Pilot 2B: Scrambled Sentence Task Pilot – Mturk Sample

Method

Participants and Design

Of all 60 participants, 85% were White, 6.7% were Black, 5% were Asian, and 3.3% were Latino. Of the 1.7% who provided text responses to the question, they identified themselves

as Central Asian/Eastern European. Economically, 6.7% identified themselves as impoverished, 15% lower class, 31.7% lower-middle class, 40% middle class, and 6.7% upper-middle class. Only one participant was excluded from the final data for failing an attentional check. On average, included participants completed the experiment in 14.26 minutes (SD = 4.41).

Procedure

The design and procedure of this study was identical to that of Pilot 2A, and the recruitment and payment procedure was identical to that of Pilot 1A.

Results and Discussion

One-way between-subjects ANOVAs were conducted on all dependent measures. A one-way ANOVA on the IPIP variety-seeking and openness measures revealed significant effects of the manipulation, $F(2, 56) = 1.85$, $p = .047$, $\eta^2 = .10$, and $F(2, 56) = 1.55$, $p = .014$,$\eta^2 = .14$, respectively. No post-hoc tests were significant for variety-seeking (see Table D for descriptive statistics). Post-hoc tests for openness using the Bonferroni adjustment found that participants in the distrust condition reported less openness ($M = 3.35$, $SD = .63$) than participants in the control condition ($M = 3.87$, $SD = .56$), $p = .02$, $d = .87$ and participants in the trust condition ($M = 3.82$, $SD = .54$), $p = .05$, $d = .80$. Tests for all measures are reported in Table 4.

Pilot 3A Impostor Detection Task Pilot – Student Sample

Other researchers have used ‘gender impostor’ paradigms in the past (Schul et al., 2004; 2008), in which answers to various gender-salient questions, such as how to change a flat tire, or what the contents of one’s purse are, are provided by some actual women, and some men are asked to give responses in which they try to impersonate a woman. Participants are then asked to rate the spontaneity of each respondent in the control condition, and to try to identify the impostors in the distrust condition. I found it necessary to adapt this manipulation in a way that

would preserve the emphasis on deception-detection while excluding the casual sexism and transphobia that might be implied by a gender impostor task. Student athletes were chosen as a substitute for women because, like gender, it is both a socially meaningful and relatively impermeable social category on a large university campus.

Method

Participants and Design

Of all 85 participants, 22.4% were male, and 77.7% were female. In terms of race and ethnicity, 72.9% were White, 5.9% were Black, 22.1% were Asian, and 15.3% were Asian, and 7.1% were Hispanic/Latino. Of those 2.4%who provided text responses to the question, they identified themselves as human, and White/Asian. Economically, 1.2% identified themselves as lower class, 10.5% lower-middle class, 45.9% middle class, 36.5% upper-middle class, and 5.9% upper class. They were 18.5 years old on average.

This pilot study was simple two-condition experiment. Participants were randomly placed in the *distrust* or *control* condition. Participants took an average of 17.85 minutes to complete the experiment (SD = 9.18).

Materials

To create materials for use in the manipulation, six volunteers were recruited; two were student athletes at the University of Connecticut, and four were non-athletes. Each of them was asked a series of questions about their everyday life and their participation in college athletics. Athletes were asked to respond truthfully, and non-athletes were asked to respond as if impersonating a student athlete. The questions were as follows.

1. What are the most important traits for a successful athlete to have?
2. What sport do you play for this school?

3. What do you enjoy most about your sport?
4. What do you like to do in your free time?
5. If you could change one rule about your sport, what would it be?
6. What is the hardest part about being a student athlete?
7. What is your major area of study?
8. What is most satisfying about your sport?

Two of the four non-athletes were selected for use in the study, so there would be equal representation of both groups.

Participants in the *distrust* condition were told that some of the respondents were imposters, and were asked to judge whether each respondent was an athlete or an impostor. Participants in the *control* condition were uninformed that some were impostors, and were only asked to rate the spontaneity of each respondent. The complete text of the manipulation is provided in Appendix L.

Measures

The dependent measures used in this pilot study were the same as the other pilot studies in this series, except the item measuring trust for police was unintentionally excluded.

Procedure

All participants were recruited online from the University of Connecticut Psychology participant pool. The study was administered electronically.

After reading the information sheet, the program randomly assigned each participant to the *distrust* or *control* condition. Participants in the distrust condition received the following instructions:

> You are about to read some statements collected from some athletes at the University of Connecticut about life as a student athlete.
>
> Here's the trick - some of these responses were collected from actual student athletes, and some of them were collected from non-athletes trying to impersonate athletes.
>
> Your task will be to carefully read the responses, and to determine which responses were given by athletes, and which ones were given by impostors.
>
> First, just carefully read each of the responses. After you have read each one, they will be shown to you again for you to make your final decisions.

Participants in the control condition received these instructions instead:

> You are about to read some statements collected from some athletes at the University of Connecticut about life as a student athlete.
>
> Your task will be to carefully read the responses, and to judge the spontaneity of each respondent.
>
> First, just carefully read each of the responses. After you have read each one, they will be shown to you again for you to make your final decisions.

Participants were given four sets of responses to the eight questions, half of which were actually from athletes and half which were not. Afterward participants in the distrust condition were told:

> Now, you are going to see the very same sets of responses. Now that you have seen them all and had a chance to think about them, indicate whether you believe the respondent is a genuine student athlete or an impostor.

and were asked to read back over the responses and for each, give a dichotomous response as to whether the individual answer was said an impostor or a genuine athlete.

For participants in the control condition, their instructions prior to rereading the answers were as follows:

> Now, you are going to see the very same sets of responses. Now that you have seen them all and had a chance to think about them, indicate how spontaneous you feel each respondent is.

They gave their spontaneity ratings on a scale of 1 to 6. Afterwards, participants completed the same dependent measures used in Pilot 1A, and same attentional checks used in Pilot 2A, and debriefing procedure

Results

T-tests were used to test for differences between conditions in each of the measures. All of these tests are reported in Table 5.

Confidence and Personal Control

Participants in the distrust condition reported lower judgmental confidence (higher uncertainty) (M = 3.35, SD = .85) than those in the control condition (M = 3.81, SD = .66), $t(62.24) = 2.52$, $p = .014$, $d = .60$.

Domain-Specific Trust

Unexpectedly, participants in the distrust condition reported more trust for their relatives (M = 4.53, SD = .71) than control participants (M = 4.11, SD = .84), $t(69) = -2.27$, $p = .026$, $d = .54$.

IPIP Variety and Openness

Contrary to expectations, participants in the distrust condition judged themselves to be less variety-seeking (M = 3.26, SD = .58) than control participants (M = 3.54, SD = .45), $t(69) = 2.29$, $p = .025$, $d = .54$, and marginally less intellectually open (M = 3.41, SD = .52) than controls(M = 3.64, SD = .53) $t(68) = 1.85$, $p = .068$, $d = .44$.

Discussion

This impostor detection task was more effective in generating effects than the semantic prime-based methods in Pilots 1 and 2. Participants in the distrust condition actually reported higher uncertainty in their judgments, as predicted, which is encouraging. Distrusting participants did, however, report greater trust for their relatives, though this could possibly indicate increased *relative* trust for people who are socially closer to oneself in response to the deceptive behavior of strangers – an interpretation supported by the lower scores on variety seeking and intellectual openness.

Pilot 3B Impostor Detection Task Pilot – Mturk Sample

Methods

Participants and Design

All participants were recruited online from Amazon.com's MTurk. Of all 70 participants, the average age was 32.18. Ethnically/racially, 71.4% were White, 8.6% were Black, 4.3% were Asian, and 2.9% were Hispanic/Latino. Economically, 1.6% identified themselves as impoverished, 19.7% lower class, 27.9% lower-middle class, 31.1 % middle class, and 19.7% upper-middle class. Fifty-five participants met inclusion criteria. On average, included participants completed the experiment in 11.4 minutes (SD = 3.54).

Procedure

The experimental procedure was the same as with the student sample above, except that participants received payment of $0.70 for their time.

Results

T-tests were performed to test for differences between conditions. See Table 6 for descriptive statistics on all measures.

Trust for Relatives and Neighbors. Distrusting MTurk workers reported significantly less trust in their relatives (M = 3.77, SD = 1.040) than control MTurk workers (M = 4.36, SD = .860), $t(53) = 2.276$, $p = .027$, $d = .62$.

They also reported marginally less trust in their neighbors (M = 3.30, SD = .837) than controls (M = 3.72, SD = .891) $t(53) = 1.8$, $p = .078$, $d = .49$.

IPIP Mistrust. Participants in the distrust condition reported greater mistrust (M = 2.55, SD = .833) than control participants (M = 2.00, SD = .798) $t(53) = 2.494$, $p = .016$, $d = .68$.

IPIP Variety Seeking. Participants in the distrust condition reported greater variety seeking (M = 3.47, SD = .726) than control participants (M = 2.99, SD = .914) $t(53) = 2.193$, $p = .033$, $d = .59$.

Interpersonal Trust and Caution. On Yamagishi and Yamagishi's (1994) measure of interpersonal trust, distrusting participants reported less trust (M = 3.30, SD = .694) than control participants (M = 3.74, SD = .895) $t(53) = 2.048$, $p = .046$, $d = .55$.

Distrusting participants also reported marginally greater caution toward others (M = 3.30, SD = .713) than control participants (M = 2.90, SD = .796) $t(53) = 1.95$, $p = .056$, $d = .53$.

Discussion

It is an encouraging sign that every one of these effects were in the predicted direction, and that there were so many of them. But why these effects when UCONN students, who are presumably more involved in discerning imposter athletes from genuine athletes, show such a different pattern? My best guess is that is has to do with the fact that MTurk workers moved through the study considerably more quickly than students did, and thus, didn't think as carefully about how they were responding to the different questions, allowing automatic processes to have a greater influence on their responses.

Pilot 4A Economic Deception Game – Student Sample

Posten and Mussweiler (2013) have used an economic deception game to create a state of distrust. This is an appealing paradigm because it actually creates a situation in which participants encounter an attempt at deception and exploitation, a situation which could realistically promote distrust in everyday life. Their procedure was adapted for use in the Qualtrics system, with English-speaking participants.

Method

Participants and Design

Of all 358 participants, 71.2% were White, 5.6% were Black, 16.5% were Asian, .8% were Native Hawaiian/Pacific Islander, and 8.7% were Hispanic/Latino. Of those 2.2% who provided text responses to the question, they identified themselves as Asian/White, Greek, Haitian, Jamaican, Mixed, South Asian, and West Indian. Economically, 1.1% identified themselves as impoverished, 2.8% lower class, 12.8% lower-middle class, 42.7% middle class, 35.8% were upper-middle class, 3.9% upper class, and .3% super rich. They were 18.9 years old on average. Inclusion criteria were the same as for the previously reported studies. There were 306 participants that met inclusion criteria. Included participants finished the experiment in 16.56 minutes (SD = 3.18).

This pilot study was a two-condition experiment. Participants were randomly assigned to the *distrust* or *trust* condition.

Measures

The dependent measures used in this pilot study were the same as Pilot Study 1A except the item measuring trust for police was unintentionally omitted.

Procedure

All participants were recruited online from the University of Connecticut Storrs campus psychology participant pool.

The study was administered via Qualtrics. Directly after reading the information sheet, participants were presented with a set of instructions describing the economic game they would be playing in detail:

> When you go to the next page, you will be connected to another participant. Together, the two of you will be deciding how to distribute a maximum of 6 virtual raffle tickets for a $20 Amazon gift card between yourselves.
>
> There are two ways the tickets can be distributed:
>
> - Each Player gets 3 tickets
>
> - 4 tickets for Player 1 and 1 ticket for Player 2
>
> You will be randomly assigned to be either Player 1 or Player 2. Each player will have a unique role in the decision-making process:
>
> - Player 1 will be able to see which option (A or B) represents which distribution, and will recommend one of the options to Player 2.
>
> - Player 2 will not be able to see which option represents which distribution, but will receive Player 1's recommendation and will make the final decision about which option will be chosen.

After reading these instructions, participants were quizzed for comprehension of the instructions. Those who failed the test were given one more chance to read the instructions and answer correctly.

Participants were then presented with a screen claiming to wait for another participant to connect. Actually, the webpage was programmed to wait eight seconds before moving on to the next screen.

The next screen was timed for three seconds, and claimed to be assigning roles to the participants. All participants were "assigned" to Player 2. The website paused for three seconds for the simulated "Player 1" to evaluate their options.

In all cases, Player 1 suggested option A, and provided a text message, "that gives each of us 3." Up to this point, both experimental conditions were identical. They diverge only in the outcomes of participant responses.

In the *trust* condition, participants who did go with Player 1's advice to pick option A received the following message in large print:

> The Selected Distribution is:
>
> You: 3 TICKETS
> Player 1: 3 TICKETS
>
> If you had gone against your partner's advice, you would have received 1 ticket, and they would have received 4.

If they picked option B, *against* Player 1's advice, they would receive this message:

> The Selected Distribution is:
>
> You: 1 TICKET
> Player 1: 4 TICKETS
>
> If you had taken Player 1's advice, you would each have been awarded 3 tickets.

Participants in the *distrust* condition were given essentially the same messages, but in reverse. They were given the egalitarian outcome (3 - 3) if they went against Player 1's advice, and were given the disadvantageous outcome (1 – 4) if they followed Player 1's advice. For a view of how this looks within the qualtrics software, with branching logic, see Appendix M.

After the manipulation phase was complete, participants were given the same measures as in the above reported pilot studies.

After dependent measures were taken, participants were asked to type an open-ended response to the question, “What do you think about your interaction partner (Player 1)?” They were also asked whether or not they thought their partner thought they were real (yes/no), and whether they believed their partner was real (yes/no). Then they were told that their partner was not real, and were asked to report on a seven-point scale how surprised they were to find this out.

Results

All dependent variables were analyzed using 2 (Manipulation: trust vs. distrust) x 2 (Belief: Believed partner real vs. did not) between-subjects ANOVAs. The latter variable was included in order to statistically control for whether or not the contrivance of the manipulation was convincing, and to determine the role of the believability of the scenario. It was possible that the economic game could only change state distrust if the participant believes they are interacting with a real person (121 reported that they believed their partner was a real person), but it was also possible that the scenario itself could affect a person’s level of distrust regardless of their belief that they were dealing with another human directly. The majority (94%) of participants took Player 1’s advice and chose option A. Results for each measure can be found in Table 7.

Confidence and Personal Control. Participants with an untrustworthy partner had less judgmental confidence/certainty (M = 3.15, SD = 1.10), than those with a trustworthy partner (M = 3.99, SD = .85) $F(1, 299) = 47.75$, $p < .001$, d = .85, and felt less personal control over life events (M = 3.39, SD = .98) than those with a trustworthy partner (M = 3.84, SD = .81) $F(1, 299) = 18.15$, $p < .001$, d = .50. Belief had no significant main or interactive effects (ds<.09).

IPIP Mistrust. As predicted, participants in the untrustworthy partner condition reported greater mistrust of others (M = 2.95, SD = .35) than those in the trustworthy partner condition (M = 2.80, SD = .32) $F(1,299) = 16$, $p < .001$, d = .45.

Matchstick Puzzles. As in the previous pilot, participants with an untrustworthy partner solved fewer of the "routine" matchstick problems (M = .90, SD = 1.04) than participants with a trustworthy partner (M = 1.18, SD = 1.20) $F(1, 299) = 6.37$, p = .012, d = .25. Participants in this condition also solved fewer matchstick problems overall (M = .99, SD = 1.25) than participants with a trustworthy partner (M = 1.30, SD = 1.38) $F(1, 299) = 6.47$, p = .011, d = .24. There was a significant interaction between belief that Player 1 was a human and whether Player 1 was a trustworthy or untrustworthy partner for the total number of matchstick problems solved $F(1, 299) = 4.11$, p = .044. There appeared to be no partner effect when participants were not convinced of their partner's realness (trustworthy partner: M = 1.17, SD = 1.29; untrustworthy partner: M = 1.09, SD = 1.32, d = .06), and a decrement in puzzles solved for participants with untrustworthy partners whom they believed to be real (trustworthy partner: M = 1.48, SD = 1.48; untrustworthy partner: M = .74, SD = .99, d = .59) .

IPIP Openness. Distrusting participants reported less intellectual openness (M = 3.43, SD = .71) than trusting participants (M = 3.65, SD = .53) $F(1, 299) = 5.69$, p = .018, d = .35.

Discussion

The economic deception game seems to have successfully influenced participants' levels of trust, moving several trust measures in the appropriate direction, with the exception of the IPIP intellectual openness measure, for which participants in the distrust condition became more intellectually conservative. Pairing with an untrustworthy partner also appears to cause participants to solve fewer matchstick puzzles, which could also indicate increased intellectual conservatism, or cautiousness. Participants paired with an untrustworthy partner were significantly less confident in the correctness of their judgments, reported lower feelings of

personal control over what happens to them, reported greater mistrust for other people generally, solved fewer matchstick puzzles, and reported less intellectual openness.

Pilot 4B Economic Deception Game – Mturk Sample

Method

Participants and Design

Of all 60 participants, 73.3% were White, 13.3% were Black, 11.7% were Asian, and 3.3% were Hispanic/Latino. Economically, 15% identified themselves as lower class, 26.7% lower-middle class, 46.7% middle class, and 11.7% upper-middle class. They were 36 years old on average. Included participants completed the experiment on average in 16.62 minutes (SD = 4.95).

This pilot study was a simple two-condition experiment. Participants were randomly placed in the *distrust* or *trust* condition.

Procedure

All participants were recruited online on MTurk. The procedure of this study was identical to that of the student sample, with the exception that this version of the study had participants competing for raffle tickets for a $5 bonus instead of a $20 gift card. Participants were also paid $0.70 for completing the study.

Results

Analyses

Inclusion criteria were the same as for the previously reported studies. Fifty-three participants met inclusion criteria. Participants were distributed fairly evenly between the four conditions – of those assigned to a trustworthy partner, 15 believed them to be real, and 13 disbelieved. Of those with an untrustworthy partner, 14 believed, and 11 disbelieved. Most

participants (86%) took Player 1's advice and chose option A. Results for all measures can be found in Table 8.

Confidence and Personal Control. Distrusting participants reported less confidence (M = 3.47, SD = 1.33) than trusting participants (M = 4.19, SD = .79) $F(1, 49) = 5.67$, $p = .021$, d = .66.

Matchstick Puzzles. A 2 (Experimental condition: trust vs. distrust) x 2 (Belief in partner: Believed partner real vs. did not) between-subjects ANOVA revealed an interaction effect on nonroutine matchstick puzzle performance, $F(1, 49) = 4.47$, $p = .04$, $\eta^2 = .08$. Among participants who believed in their partner, distrusting participants appeared to have solved more of the puzzles (M = .27, SD = .65), than trusting participants (none – M = 0, SD = 0) , d = .59. For those who did not believe their partner was a real person, the opposite pattern emerged – distrusting participants solved no puzzles while trusting participants solved some, though less than one on average (M = .13, SD = .35), d = .52.

IPIP Variety. Participants who believed in their partner reported being more variety seeking (M = 3.72, SD = .79) than participants who did not believe in their partner (M = 3.17, SD = .77) $F(1, 49) = 6.03$, $p = .018$, d = .71.

No other effects were significant for any measures.

Discussion

In spite of a considerably smaller sample, this iteration of the economic deception game still managed to replicate significant effects in terms of confidence in the correctness of one's judgments, and the ability to solve complex puzzles. Regardless of belief in their partner, participants in the distrust condition reported feeling less confident in their own thoughts and perceptions. This distrust manipulation also had an effect on participants' ability to solve

counterintuitive puzzles, though the direction of that effect was subject to whether the participant believed their partner was a real person. Encountering an untrustworthy partner made people solve more puzzles, provided they believed that partner was a real person. This is in line with previous findings (Schul et al., 2008), and suggests that is will be important, when using this type of paradigm in an online setting, to evaluate whether or not participants believe they are dealing with a real human being.

General Discussion of Pilot Studies

General Overview of Pilot Results

Across populations and methods of delivery, semantic primes were not found to be a reliable means of manipulating reported levels of trust, distrust, caution, or performance on cognitive tasks.

Of all the priming-driven pilot studies, the most compelling was the word search study with a student sample, which still failed to produce effects of encouraging size, all of which were in the opposite of the predicted direction. The same method produced no effects with an MTurk sample. The scrambled sentence task failed entirely in both populations to produce changes in the dependent measures.

A recent meta-analysis found that direct behavioral primes have their strongest effect when the target behavior has high value to the participant, and that the effect last longer when the prime activates a goal (Weingarten, Chen, McAdams et al., 2016). Although the analysis was restricted to direct behavioral primes, and the experiments here have attitudinal and indirect targets (e.g. distrust primes and similarity judgments), it is worth pointing out that the target tasks were unlikely to have a high value to participants and did not activate any goals. In that respect, these results were congruent with the larger priming literature.

Fortunately, the manipulations that simulated actual distrust-relevant social situations worked considerably better. Surprisingly, the impostor detection task was very effective at reducing trust, increasing caution, and even increasing self-reported variety-seeking, all movements in the predicted direction for MTurk workers – all this despite being specifically designed for University of Connecticut students. The UCONN students did show effects of condition in this manipulation on measures of confidence, trust, variety-seeking and intellectual openness, but only the effect on confidence was in the predicted direction.

In another surprising turn of events, MTurk workers did not seem to be as impacted by the economic deception game as UCONN students were, despite the manipulation being designed with their population in mind. MTurk workers in the distrust condition were feeling more general mistrust according to the IPIP scale, but also felt more trusting of police. Meanwhile, students in the distrust condition of the economic deception game reported lower levels of confidence and control, and lower levels of trust on a variety of measures. Also, when they believed their interaction partner was human, they solved more counterintuitive matchstick puzzles in the distrust condition than in the trust condition, which is in line, though not perfectly so with Schul et al.'s (2007) finding that distrusting participants performed better on the counter-intuitive set, but worse on the intuitive set than did trusting participants.

Differences between Methods

Lack of priming effects. Two methods of semantic priming in two different populations failed to provide compelling evidence for the effectiveness of distrust primes. In recent years, social psychological priming has been embroiled in controversy, with some large-scale studies failing to replicate classic effects (Harris et al., 2013; Pashler, Coburn, & Harris, 2012; Pashler et al., 2013; Shanks, Newell, Lee, Balakrishnan, Ekelund, Cenac, Kavvadia, & Moore, 2013).

However, semantic primes *have* been successfully used to induce distrust in the past (Posten & Mussweiler, 2013; Mayer & Mussweiler, 2011; Friesen & Sinclair, 2011), so why not here?

Language and culture. One of the most pressing reasons to conduct pilot studies on the semantic priming of distrust in the first place was that all of the other studies that had successfully used the method were conducted outside the US, and usually in the German language (Posten & Mussweiler, 2013; Mayer & Mussweiler, 2011; Friesen & Sinclair (2011) was conducted in Canada, in English). It is possible that the words used in the German studies had somewhat different connotations or associations than their translations do in American English. It is also possible that people in different cultures react to the concept of distrust in different ways. For instance, people in societies with more social assurances find interpersonal trust to be a less important factor in their interactions with others, because they can rely on their social and cultural institutions to ensure that it is in the other person's best interest to behave in a sufficiently trustworthy way, regardless of that person's feeling toward them (Yamagishi & Yamagishi, 1994).

Increased trust? Students who were given distrust primes embedded in a word search were actually *more* trusting on some measures. The first explanation that might come to mind is that participants were reacting against the manipulation and working to re-establish their ability to trust others. This explanation seems unlikely, however, because other methods of inducing distrust, such as the impostor detection task and the economic deception game produced even more effects in the predicted direction.

A more likely explanation is that this effect has to do with the semantic characteristics of distrust-relevant terms. In the English language, distrust is semantically a negation of trust. It is difficult to describe distrust without referencing the concept of trust. Prime words for distrust

inevitably included such entries as *dishonest* or *uncertain* – negations of trust-evoking words. Even words that do not include trust-words, such as *fake* or *suspicion* are difficult to comprehend without contrasting them to concepts like *authentic* or *security*. If this is the case, then semantic distrust primes may be inadvertently priming for trust at the same time to some degree.

Non-priming methods. The imposter-detection task and economic deception game likely worked better than the semantic primes because they created a situation in which participants were required to engage in some of the mental tactics that are necessary in real life distrust-relevant situations – looking for signs of deceit in the imposter-detection task, and being betrayed/successfully avoiding betrayal in the economic deception game. In short, they compelled people to *be* distrusting, rather than make them think about distrust.

In fact, these methods may have succeeded for the same reason that priming failed – that is, the fact that they promoted effortful, conscious distrust may have been key to their success. If distrust is active rather than passive, manipulations that promote active distrust may be necessary to achieve distrust effects. The fact that such an overwhelming majority of participants in the economic deception game went along with Player 1's advice further supports the idea that trust is the default "mode," and that distrust requires some effort.

Differences between Sample Populations

Impostor detection. Why did MTurk workers respond so much better to the impostor detection task than UCONN students? After all, the manipulation was designed specifically for students at the Storrs campus. It is surprising that MTurk workers, who presumably have no personal connection with UCONN athletics, would show greater distrust effects than the students, for whom the ability to tell genuine athletes from fakers is more personally relevant.

One possible explanation is that students' social proximity to the issue made it psychologically necessary for them to personally reaffirm that they could trust the people close to them, even if some students are deceitful. Because these were students who lived and learned among other UCONN students, it may be important to them to believe they can trust the people around them. The fact that students in the distrust condition did report less confidence in the correctness of their opinions *and then* reported greater trust in their relatives and other individuals, and less variety-seeking and intellectual openness suggests that they could be responding to a threat to their personal confidence by bolstering their trust in their own group members and avoiding outside influences.

It is also possible that the effects of the distrust manipulation are short-lived. Students took an average of six minutes longer to complete the experiment than MTurk workers, so they may have had more time post-manipulation to overcompensate for their uncomfortable feelings of distrust.

Economic deception. Just as MTurk workers responded better to the manipulation designed for university students, the students responded better to the manipulation that was implemented with MTurk workers in mind – the economic deception game. Where MTurk workers showed slightly-*increased* trust for police, and increased mistrust on the IPIP mistrust scale, students reported reduced confidence and control, and reduced trust on a majority of the measures we used.

One reason for this may be a difference in motivation. Although a $5 bonus for such a short task is almost unheard of on MTurk, it may not be as motivating as the $20 gift card that university students were competing for. Additionally, MTurk workers may not have felt confident that the bonus offer was legitimate. Researchers have obtained a bad reputation among

MTurkers for taking too long to pay, or not paying at all (Jacques & Kristensson, 2013). Bonuses in particular may be considered by some to be unreliable, as they must be given out manually, and the process is unintuitive.

The manipulation also may have seemed more novel to students, and thus, more worthy of their attention. It is relatively common that MTurk workers are asked to interact with one another in some mediated way. They may have just gone through without giving the game much thought. Additionally, MTurkers were purportedly interacting with some random person in the world, while students were led to believe that they were interacting with a fellow UCONN student, possibly even a classmate. This may have given the betrayal more of a "sting" for the students, impacting their responses on the subsequent measures.

Distrust or Uncertainty?

It is possible that the outcomes observed in this series of pilot studies is due to the manipulation of participants' feelings of certainty, and not distrust *per se*. These experiments did find some effect of the distrust manipulations on reported personal control and certainty (Marsh & Weary, 1994) among students in the impostor detection and economic deception studies. At this point, it is even difficult to distinguish between uncertainty and distrust. Uncertainty and lack of personal control has been linked with more conservative, deliberative, and diagnostic processing of information (Marsh & Weary, 1994; Pittman & D'Agostino, 1985; Pittman & Pittman, 1980; Weary et al., 1993), which seems to be very similar to some of the outcomes associated with distrust – reduced stereotyping (Posten & Mussweiler, 2013), activation of more social categories in response to a complex stimulus (Friesen & Sinclair, 2011), and higher performance with complex puzzles (Schul et al., 2008).

Control motivation and uncertainty do not, however, seem to explain the apparent preference for alternatives and simultaneous activation of conflicting concepts (Friesen & Sinclair, 2011; Schul et al., 2004), or increased creativity (Mayer & Mussweiler, 2011) also included among the non-routine cognitive patterns associated with distrust. The question also remains whether experimentally manipulated distrust is conceptually similar enough to be equated with more stable concepts of distrust, such as distrust of institutions or outgroup members.

Conclusion

The most important thing to learn from these pilot studies is that, regardless of sample population, online priming manipulations have not been shown to be a reliable means of inducing a state of distrust, at least according to the variety of measures used here. Although both imposter-detection and economic deception manipulations have been fairly effective, they do not always have the same effects across populations. Moving forward, only the non-priming methods will be used to induce distrust.

In the upcoming studies, both the imposter detection and economic deception procedures will be used for both populations, though the majority of studies performed on students will make use of the economic deception task. While the effects of these manipulations are now known for these pilot measures, they of course remain unknown for the measures being used in the coming studies, so some variety in the manipulations used is desirable.

Chapter 8

Experiment 1

This study used one of the manipulation techniques that worked relatively well in pilot studies, the impostor detection task, in order to investigate how distrust changes the way that statements are interpreted. In particular, it addressed the type of asymmetrical criticism that has frustrated advocates of science-based medicine for some time. Many people who are highly critical to the point of rejecting established medical interventions such as vaccines or antibiotics for their children, also uncritically accept CAM treatments for which there is, at best, dubious evidence. The existence of this unbalanced criticism implies that distrust does not lead to a generalized resistance to persuasion, but instead leads to a systematic bias toward criticizing a default, or established position.

Previous research on distrust has demonstrated that distrust, as a state, appears to facilitate a readiness to identify information that contrasts with what they are being exposed to currently – for example, by identifying contrasting adjectives more quickly, or judging members of an outgroup less stereotypically (Schul et al., 2004; Schul et al., 2008; Posten & Mussweiler, 2013). This seems to imply that distrusting people will find a statement to be comparatively less likely when it agrees with an earlier, established statement – reading the first statement should prompt the search for contrasting information, making the eventual discovery of an agreeing statement feel less "correct" because it matches less closely with the information they were searching for, whereas trusting people may find the agreeing statements more likely to be true, because the statement seems familiar to them.

In this study, participants were placed either in a situation where they needed to judge whether or not someone was lying to them (distrust), or in which they needed only to judge their

spontaneity (control), and were then exposed to two sets of statements. The first set included a variety of trivia-type statements, some of which were true, some false. The second set also consisted of trivia-type statements, each of which either agreed with, contradicted, or were irrelevant to one of the statements in the first set. It was predicted that participants in the distrust condition would be searching for information that contrasts with the first set of statements, making the later contradictory statements seem more likely to be true, due to this increased availability of ideas that contrast with the statements in the first set. Some of the questionnaire measures from the pilot studies were included in the interest of exploring whether or not self-report distrust measures also influenced belief in three statement types.

Method

Participants

Sixty-two University of Connecticut students followed the link to the Qualtrics survey. Of those who reported their gender and ethnic background, and 67.7% female; 66.1% were White, 4.8% Black, 17.7% Asian, 3.2% Native Hawaiian/Pacific Islander, and 6.5% Hispanic/Latino. One student identified as a member of the Blackfoot Tribe, and one as Sri Lankan. Economically, 1.6% identified as lower class, 11.3% lower-middle class, 45.2% middle class, 35.5% upper-middle class, and 3.2% upper class. The students were 19 years old on average. Students were assigned randomly to the distrust or control condition.

Materials and Procedure

Participants were recruited through the psychology department's participant pool website. Upon signing up, they were presented with a link to the Qualtrics survey that contained the study.

After reading the information sheet for the study, participants were taken through the impostor detection version of a judgment task or a control version of the judgment task.

Impostor detection judgment task. The athlete impostor detection task piloted in the previous chapter was used to induce a state of distrust in participants in the distrust condition. Participants were presented with four respondents' answers to eight questions about being a student athlete. In the distrust condition, participants were informed that two of the four respondents were genuinely student athletes at the University of Connecticut, and two were non-athletes doing their best impersonation of a student athlete. They were asked to read each of these responses and for each, give a dichotomous response indicating whether they thought each respondent was genuine or an impostor.

Control judgment task. In the control condition, participants were presented with the same responses about being a student athlete, but participants were not told that some of the responses were actually from non-athletes. Instead, participants were asked to read the responses and rate the spontaneity of the response on a six-point scale (as in Schul, Mayo, Burnstein, 2004).

Statements. Once the manipulation was completed, students completed an unrelated second task in which they were asked to rate the plausibility of various statements. Thirty statements were facts taken from a cultural literacy test (adapted from Julka, 1993) based on a book about cultural literacy covering basic facts of history, literature, and science that literate Americans should know (Hirsh, 1988). The statements were split into two 15-statement halves. The primary dependent measure was the responses to the second set of 15 statements, with these target statements belonging to one of three categories. Five statements were directly *contradictory* to a statement presented in the first half, five were directly in *agreement* with a

statement presented in the first half, and the five statements in the third condition were *irrelevant* to all previous statements. To create contradictory statements, one detail of a statement in set one was changed, and the statement rephrased in order to make the contradiction less obvious. Statements were presented in the same semi-randomized order for all participants. To eliminate the confound of participants only being exposed to *false* statements that contradict an earlier statement, two contradictory target statements were true (and thus, the corresponding set 1 statement false), and three of these target statements were false. The two halves were presented to the students as one continuous set. Eight of the 15 statements in the first set, and six of the 15 target statements were factually true (See Appendix N).

Each statement was presented to participants on its own webpage, to prevent them from looking back or ahead at other statements while responding. After reading the statement, the participant was asked, "please estimate your confidence in the truth of the following statement." Plausibility ratings were recorded by having the participants move a cursor between 0 and 100 on a slide located just below the statement, with the cursor's starting location automatically appearing at 50 before participants responded. (See Appendix O for a visual depiction). The focus of my analyses were responses to the second set of statements, broken into three groups of five, and averaged, yielding one score for each type – *agreeing, contradicting,* and *irrelevant* statements. Finally, the participants completed the following self-report measures.

Uncertainty and perceived control. The items that were most sensitive to the manipulations in the pilot studies were included for assessing degree of confidence and certainty in one's judgments and degree of perceived control over outcomes. Two single-item measures used in research on chronic uncertainty and perceived control deprivation (Marsh & Weary, 1989; Weary, Marsh, Gleicher, & Edwards, 1993) were adapted for assessing transient feelings

of undermined confidence and certainty and transient feelings of undermined feelings of personal control. These items were included as manipulation checks to assess the degree to which a person is uncertain or is confident in the correctness of their opinions, judgments, and decisions, and how much they feel that they lack or have sufficient personal control over the things that happen to them.

Medical mistrust. The Revised Health Care System Distrust Scale (Shea, Micco, Dean, McMurphy, Schwartz, & Armstrong, 2008; Appendix P) was used as one measure of real-world distrust. This scale measures two dimensions of mistrust for the health care system – how competent participants feel the health care system is (MMC; $\alpha = .76$), and how much its values align with their own (MMV; $\alpha = .68$). These dimensions correspond rather well with the present notion that distrust involves a conflict of interest between parties. Recent studies have found that African-Americans indicate higher distrust on this scale, particularly on the values dimension (Armstrong et al., 2013), and this relationship appears to be moderated by the amount of racial discrimination a person experiences (Shoff & Yang, 2012).

This scale was used as the primary dispositional measurement of distrust because it conveys distrust in an explicit, real-world entity. Most measures of trust and distrust ask participants to report their agreement with general statements about how trusting they are, and how selfish or untrustworthy others are (Goldberg et al., 2006; Yamagishi & Yamagishi, 1994). Such measures can be expected to reflect a person's opinion of how much they trust others generally, but may not reflect the way they respond to entities that ask for their trust. The Revised Health Care System Distrust Scale (Shea, Micco, Dean, McMurphy, Schwartz, & Armstrong, 2008) evaluates actual beliefs about a major social institution. Instead of measuring distrust as a kind of misanthropy, this scale measures it as in terms of their beliefs about a

specific entity's competence and moral integrity. It was included as a means to conduct exploratory analyses concerning how this form of dispositional distrust relates to the manipulated distrust found in laboratory studies.

Trust For Specific Other Persons and Institutions. As in the pilot studies, participants were asked to report how much they trusted their neighbors, relatives, self, science, police, and religion. Additionally, they were asked to report how much they trusted their government. Each of these measures was a single item.

IPIP Variety and Openness. IPIP scales for variety-seeking (α = .83) and intellectual openness (α = .66; Goldberg et al., 2006) were also used again for this study, because they often were affected by this manipulation in pilot studies, and serve as a way of measuring people's self-perceived open-mindedness, which may relate to their level of cognitive flexibility. As such, these measures were expected to actually increase under distrust.

Additional items. At the end of the study, participants completed some exploratory items as part of a separate research project examining what dispositional and attitudinal factors might be associated with attitudes toward vaccines. These measures were included at the end of all subsequent studies.

Social dominance orientation. The four-item Short Social Dominance Orientation (SSDO; α = .791) scale was given to participants to measure their favorability toward group-based inequality (Pratto et al., 2013).

Right wing authoritarianism. The RWA scale was used as a measure of authoritarianism (α = .931; Altemeyer, 2006).

Abortion, vaccines and politics. Participants completed a variety of items assessing their attitudes toward abortion rights and childhood vaccines. They also completed a measure assessing their political ideology (far-left liberal to far-right conservative).

Results

Data Preparation

Interpretation of missing plausibility ratings. In the Qualtrics data output for plausibility ratings, about 63% of participant data contained seemingly randomly-distributed missing data, but only on the plausibility ratings and not the other measures. It did not seem likely that a participant would occasionally decline to rate the plausibility of a statement, while rating many others, and completing the rest of the study without incident. A closer examination of the display of the "slider" input device on the webpage for assessing plausibility makes it apparent that the slider had a default position at the midpoint (50). Thus it seems most likely that participants with the occasional blank space in their statement ratings meant to give it a rating of 50, rather than intentionally skipping the item. Thus, for all participants who had missing data *only* on plausibility responses, their missing ratings were assumed to be 50 in the primary analyses. Analyses were also conducted eliminating missing responses; those results did not differ substantially from the primary analyses. See Appendix O for pictures of the sliders.

Manipulated Distrust

Likelihood ratings. A 2 (Judgment type: Distrust vs. Control) x 3(Statement type: Agreeing vs. Contradictory vs. Irrelevant) mixed analysis of variance, with statement type as the within subjects variable, revealed a main effect of statement type, $F(2,120) = 12.31$, $p < .001$, $\eta_p^2 = .17$. Post-hoc tests of differences between statement types revealed that contradictory statements that contradicted a prior statement were viewed as significantly less plausible (M =

49.46, *SD* = 12.19) than agreeing (*M* = 59.62, *SD* = 13.25) $t(63) = 4.876$, $p < .001$, $d = .80$, and irrelevant statement types (*M* = 58.57, *SD* = 17.20) $t(63) = 3.72$, $p < .001$, $d = .61$ (familywise alpha = .0167). Contrary to predictions, there was no significant effect of judgment type, and no interaction between judgment type and statement type.

In order to determine the role that the truth or falseness of a statement played in participants' likelihood judgments, a 2(Judgment type: Distrust vs. Control) x 2(Statement type: Agreeing vs. Contradictory) x 2(Veracity: True vs. False) mixed ANOVA was conducted. Irrelevant statements were excluded from this analysis because they were all false. Truth value did have a main effect, with true statements judged to be more likely (*M* = 64.82, *SD* = 19.42) than false statements (*M* = 44.85, *SD* = 18.36) $F(1, 60) = 49.84$, $p < .001$, $\eta_p^2 = .45$. The main effects for Statement type remained significant, with agreeing statements judged more likely (*M* = 58.09, *SD* = 18.89) than contradicting statements (*M* = 51.59, *SD* = 18.89) $F(1, 60) = 8.81$, $p = .004$, $\eta_p^2 = .13$. There were no significant interactions between the variables on plausibility ratings.

Difference scores. Although the plausibility of the statements themselves may not have responded to judgment type, it was possible that the 'change' in plausibility from a previously-related statement could be affected. If an apparently likely statement had been previously read under distrust, the comparative likelihood of the new, contradictory statement could possibly be increased compared to what it would have been otherwise, and yet still have a low likelihood judgment. Similarly, apparently unlikely statements' contradictions may have simply been rated more likely than they would be otherwise, but when raw scores were combined, this difference would be lost. Thus an analysis that takes into account the judged plausibility of the

corresponding comparison statements in set 1 could be more sensitive at detecting distrust effects.

The difference between each of the target statements and their corresponding statement in the first half of the set was taken by subtracting the rating of the earlier statement from the later target statement. Then, the average change was calculated for each statement type, with more positive numbers meaning increased plausibility of the target statements relative to its pair. Irrelevant statements, by definition, had no corresponding statement, so the average of the first-half irrelevant statements' ratings was subtracted from the average rating of the later target irrelevant statements. Irrelevant statements were included in this manner merely to provide control for time, providing a baseline for how much statements' likelihood changed generally from set 1 to set 2. All irrelevant statements were factually false.

A 2(Judgment type: Distrust vs. Control) x 3(Statement type: Agreeing vs. Contradicting vs. Irrelevant) mixed ANOVA on the difference scores, with statement type as the within subjects variable, again revealed a main effect of statement type $F(2,120) = 6.829$, $p = .002$, $\eta_p^2 = .102$. Post-hoc comparison by statement type found that contradictory target statements were rated as less plausible than the statements in set 1 that they disagreed with ($M = -6.95$, $SD = 18.17$), compared to irrelevant (control) statements presented in the second half of the statement set ($M = 4.92$, $SD = 19.39$), $t(63) = 3.562$, $p = .001$, $d = .63$. Neither were significantly different from agreeing statements ($M = -.75$, $SD = 16.56$).

Distrust of the Medical System

To test the exploratory hypothesis that dispositional differences in trust might interact with statement type in a way analogous to what was predicted for the experimental manipulation, analyses were conducted replacing the categorical manipulated distrust variable with a

continuous measure of dispositional mistrust. This analysis also took into account the actual truthfulness of the statements; a hierarchical linear model was used to test the effects as it required analyzing the data at the level of individual statements rather than relying on composite measures as in the primary analyses reported above. Manipulated distrust did not have a significant effect on reported medical mistrust, either in terms of the values-based factor $t(58) = -.25$, $p = .80$, $d = -.06$, or the competence-based factor $t(58) = -.09$, $p = .93$, $d = -.02$.

Model theory and alignment. Because the three statement types were nested within participants, a multilevel analysis was conducted on set 2 statements, with statement type as the level-1 variable, and medical mistrust as the level-2 variables (Table 11). Statements type was further divided into true and false statements, because both the agreeing and contradictory groups contained statements that were true and false. Irrelevant statements were all false in the critical set of statements. The criterion variable was the average likelihood rating of the type of target statement (agreeing, contradictory, or irrelevant). MMC and MMV were grand-mean centered, which will be described in greater detail in the "centering" section.

Random effects. In modeling these relationships, the intercept was allowed to vary freely, as different individuals may generally tend to judge claims to be more or less plausible. The statement-type slopes, on the other hand, were held fixed, because individual differences in statement-type-based plausibility differences were not expected to be meaningful beyond what could be explained by dispositional distrust.

Presentation of the statistical model. The first model was estimated including all five statement types, and both medical mistrust factors – competence-based medical mistrust (MMC) and values-based medical mistrust (MMV). The intercept (β_{0j}) represents the estimated plausibility rating for irrelevant-false statements.

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{Agreeing-True}) + \beta_{2j}(\text{Agreeing-False}) + \beta_{3j}(\text{Contradictory-True}) + \beta_{4j}(\text{Contradictory-False}) + r_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{MMC}) + \gamma_{02}(\text{MMV})$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\text{MMC}) + \gamma_{12}(\text{MMV})$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}(\text{MMC}) + \gamma_{22}(\text{MMV})$$

$$\beta_{3j} = \gamma_{20} + \gamma_{21}(\text{MMC}) + \gamma_{22}(\text{MMV})$$

$$\beta_{4j} = \gamma_{20} + \gamma_{21}(\text{MMC}) + \gamma_{22}(\text{MMV})$$

Because MMC was not found to significantly add to the model, it was removed as a predictor, resulting in the following model:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(\text{Agreeing-True}) + \beta_{2j}(\text{Agreeing-False}) + \beta_{3j}(\text{Contradictory-True}) + \beta_{4j}(\text{Contradictory-False}) + r_{ij}$$

$$\beta_{0j} = \gamma_{00} + \gamma_{01}(\text{MMV})$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}(\text{MMV})$$

$$\beta_{2j} = \gamma_{20} + \gamma_{21}(\text{MMV})$$

$$\beta_{3j} = \gamma_{20} + \gamma_{21}(\text{MMV})$$

$$\beta_{4j} = \gamma_{20} + \gamma_{21}(\text{MMV})$$

Centering. The two factors of the revised medical mistrust scale, MMV and MMC, were both grand-mean centered in the multilevel models, because a value of zero is not otherwise helpful in interpreting the models – minimal score on the scale does not necessarily indicate a complete lack of mistrust, so it is more useful to frame the expression of the model in terms of distance from the mean. The intercept in the model therefore represents the predicted statement

plausibility for a person with the average amount of medical mistrust judging a false statement that is irrelevant to the other statements read.

Software and parameter estimation. The models in this analysis were estimated using full maximum likelihood in HLM for Windows (Raudenbush, Bryk, Congdon, 2013).

Assumptions and descriptive statistics. All variables in the model appeared to be normally distributed. Means for both medical mistrust factors were near the midpoint (MMC: M = 3.04, SD = .65; MMV: M = 2.75, SD = .64). Means and standard deviations of plausibility judgments by statement type can be found in Tables 9 and 10. The intraclass correlation (ICC) for a null model of statement plausibility was .07.

Model building. Adding statement type to the null model (the random effects model) significantly reduced the deviance, Akaike information criterion (AIC), and Bayesian information criterion (BIC), while adding four more parameters to the model. In this random effects model, contradictory-true statements were not found to be significantly different in plausibility from irrelevant statements, while agreeing-true, agreeing-false statements, and contradictory-false statements were significantly different from ratings of irrelevant statements. As expected, agreeing-true statements were estimated to be 8.49 points *more* plausible, agreeing-false statements were 9.78 points *less* plausible, and contradictory-false statements were 18.21 points *less* plausible than irrelevant statements. All models in this analysis are detailed in Table 11.

Full contextual model. The full contextual model, which added both MMV and MMC as level-2 variables, estimated each of the statement-type intercepts to be virtually identical to those in the random effects model. Because there were no significant effects of MMC, the model

excluding it will be the one discussed here. The parameter estimations of the full contextual model can be found in Table 11.

Model excluding MMC. Removing MMC from the model reduced the number of parameters from 17 to 12, while only increasing the deviance from 2588.42 to 2593.23, and resulting in lower AIC and BIC compared to the full contextual model.

False contradictory statements viewed by participants with average MMV were estimated to be judged as significantly less plausible than irrelevant statements, by 18.21 points, while true contradictory statements did not differ in likelihood from irrelevant statements among participants with the average MMV. Both true and false contradictory statement types had significant negative MMV slopes, which estimated that for each point of increased mistrust above (or below) the mean, a true, but contradictory statement would be judged 10.42 points less (or more, if below the mean) likely than an irrelevant statement would be at the same value of MMV, and a false, contradictory statement would be judged less likely by 11.79 points. These negative slopes suggest that medical distrust did not have the hypothesized effect of increasing the perceived likelihood of contradictory statements (see Figure 1 for a visual representation).

On the other hand, other patterns were somewhat more consistent with my general theory of distrust. The MMV slope for irrelevant statements was 9.84, indicating that greater medical distrust was associated with higher perceived plausibility of these false statements relative to individuals with less distrust. False agreeing statements did *not* have significantly different MMV slopes from the irrelevant statements; thus greater distrust was associated with a higher judged plausibility of these statements (relative to more trust folks) as well. True agreeing statements had a negative MMV slope of 8.58, almost entirely cancelling-out the positive slope

found for irrelevant statements, and implying that MMV does not have an effect on the judged likelihood of true agreeing statements.

Initial Statements. The first half of the comparison statements – the set 1 items that are being contradicted or agreed with by the target statements in this study – were approximately half true and half false. Set 1 statements were analyzed in order to determine whether the impostor detection task or medical mistrust had an effect on participants' impressions about statements for which these earlier statements that were yet to be supported or contradicted. Ratings of the true comparison statements in set 1 were averaged, as were the false comparison statements and a mixed effects analysis with truth value as the within-subjects variable, and values-based medical mistrust (MMV; higher values indicate more mistrust) as a continuous variable revealed that true statements from set 1 (M = 60.83, SD = 14.07) were judged to be significantly more plausible than the false statements were (M = 52.20, SD = 14.48), $F(1, 58) = 8.04$, $p = .006$, $d = .60$. MMV was not found to have a main effect on statement plausibility of set 1 statements, $F(1, 58) = 1.26$, $p = .27$, $\eta^2 = .08$, but plausibility was significantly affected by a MMV x truth value interaction, $F(1, 58) = 5.04$, $p = .029$, $\eta_p^2 = .08$.

Decomposition of the interaction found that there was a significant slope in the false statements condition. Consistent with distrust hypotheses, higher levels of values-based medical mistrust were associated with increased ratings of the plausibility of false statements, $B = 6.89$, $p = .017$; medical mistrust was not associated with differences in the plausibility of true statements, $B = -2.85$, $p = .319$.

Discussion

This study did not find the predicted effect of the impostor-detection task on the likelihood judgements of statements that contradict previous statements. In fact, the impostor

detection task was not found to have any effect on the judged likelihood of any of the statements. Thus, I did not find support for the notion that inducing distrust in Americans makes contradictions easier to believe.

Although manipulated distrust was not predictive of likelihood judgments, the dispositional measure MMV was. Those with more values-based distrust of the medical system saw only irrelevant statements (all of which were false) and agreeing-false statements as more plausible than did those who were more trusting. This seems to suggest that people high in dispositional distrust find novel ideas that do not contradict available knowledge more believable than do people who are lower in dispositional distrust. So, I have not found support for the notion that distrust promotes the belief of ideas that contradict the status quo, but the data do seem to suggest that dispositional distrust may promote greater receptivity to alternatives.

The lack of an effect of manipulated distrust may be due to a shortcoming of the impostor detection task, or it could also be possible that dispositional distrust of an institution is qualitatively distinct from prompted suspicion or skepticism of a particular person's honesty or trustworthiness. This appears likely, at least as far as the imposter detection task goes, especially in light of the finding that medical mistrust did not respond to the manipulation.

This seems to conflict with Schul et al.'s (2004) finding that participants were faster to identify target adjectives that contradicted prime adjectives. If contradictory information is more easily activated, it ought to be easier to believe. Again it seems possible that there are more than one construct at work within what we have been calling "distrust." The next study uses a different technique for manipulating distrust, and provides a different context for processing contrasting information.

Chapter 9

Experiment 2

Whereas Study 1 examined how distrust affects the judged likelihood of a statement, this study looks at how distrust influences the perceived credibility and competence of a given communicator. Recent research has shown that distrust can reduce stereotyping in trait judgments of individuals by causing the perceiver to focus on how the individual is different from the stereotype, rather than similar to it (Posten & Mussweiler, 2013), rather than assimilating that individual to the stereotypes of their salient group (Devine, 1989; Dijkterhuis & Bargh, 2001; Kawakami, Young, & Dovidio, 2002; Winter & Uleman, 1984). This effect has interesting implications for source credibility. If distrust makes individual women seem less like the stereotypical woman, individual Turks less like stereotypical Turks, and individual overweight people less like the stereotypical overweight person, then it could potentially also make experts seem less expert-like, and quacks seem less quack-like. If distrust makes group members seem less representative of their group, it could make individual medical experts' advice seem less credible, and thus easier to ignore. By the same principle, for someone who belongs to a group generally *not* considered to be credible, contrasting the individual with his or her group would make that individual *more* credible.

This effect should be reduced, or even eliminated, in circumstances where members of other, contrasting groups are also present – if an expert and a quack are presented next to each other, for instance, the most available contrasts would be those between the two individuals, not those differences between the individuals and their groups. In this context, a distrusting person should generally draw conclusions about the credibility of an expert versus a quack in the same way that a trusting person would.

In this study, participants were paired with either a trustworthy or untrustworthy partner in an economic deception game, and viewed photos and descriptions of a healthcare expert, a healthcare related non-expert, a person with no relationship to health care, and an obvious quack - an energy healer trained in vibrational therapy. They were then asked to rate each of them on a variety of traits relevant to their credibility, knowledge, and skill in providing care. They were also asked to report how good this person would be at giving medical advice to their friends. In another condition, participants saw all four characters together, and were asked to provide the same feedback for all four on the same page. When presented separately, it was predicted that the characters will be judged more moderately by participants in the distrust condition on all traits, leading to fewer differences among the characters than occurs for trusting participants. The opposite pattern of character differences is predicted when all characters are presented together to distrusting individuals (i.e. ratings of the different characters should diverge from one another, to a greater degree than they do for trusting individuals). Thus, the key prediction was that presentation (alone or together) would interact with partner (trustworthy or untrustworthy) on trait judgments of each of the key characters.

Method

Participants

One hundred thirty-seven men, 133 women, and one transgender person were recruited through MTurk, and were paid $1.50 for their participation. Of those who provided racial demographic information, 80% were White, 8% were Black, 6% were Asian, and 4% were Hispanic/Latino/a. Two percent considered themselves impoverished, 20% middle class, 23% lower middle class, 43% middle class, 11% upper middle class, and appropriately, 1% identified as economic elite. They were 37 years old on average.

Materials

Inducing distrust. This study used the economic deception game detailed in the pilot studies. Participants were randomly assigned to the distrust or control condition of the game.

Characters. For each character, a picture and a brief biographical sketch were provided, stating their credentials, their position, the services they provide, and some of their personal interests (see Appendix T). Characters consisted of a medical doctor (expert), a hospital administrator (related non-expert), a music teacher (non-related non-expert), and an intuitive healer (quack).

Trait ratings. A list of 16 traits were presented to participants, including eight traits relevant to health care – *intelligent*, *caring*, *competent*, *professional*, *skilled*, *knowledgeable about health*, *knowledgeable about the body*, and *capable of giving health advice* – and eight generic traits provided as filler– *Excitable*, *athletic*, *outgoing*, *shy*, *knowledgeable about ethics*, *knowledgeable about the mind*, and *capable of giving life advice*. Participants evaluated each character's possession of the traits on a scale ranging from 1 (not at all) to 5 (very much). Trait ratings for medically-relevant traits and medically-irrelevant traits were averaged for each participant, with higher scores indicating the character exemplifies the trait more.

Memory Check. Immediately after the trait ratings were completed, participants were briefly quizzed on their memory of the characters' biographies. This was included in an effort to identify which participants were paying close attention to the task and which were not.

Personal use of CAM treatments. Because some participants may be biased in favor of pseudoscientific health care treatments independently of the context of this study, they were asked about their use of complementary and alternative medicine (CAM) products and services, such as (non-prescribed) dietary supplements, homeopathic and herbal medicines, faith healing,

reiki, acupuncture, chiropractic, and reflexology (Appendix U). During analysis, retail products were separated from services administered by practitioners before being totaled. These ancillary measures were primarily intended to serve as controls for personal bias against mainstream medicine and in favor of the quack character.

Trust/distrust measures. The measures of confidence and personal control from the pilot studies were again used here. Participants were also asked to report their trust of the police, their neighbors, relatives, self, science, religion, and government, to be used for a different research question involving the influence of distrust on the choice to vaccinate. Distrust of the health care system was again evaluated via the revised health care system distrust scale (Shea et al., 2008). This measure was included in order to examine whether people respond differently to their partner based on their expectations of fair treatment and honesty in the real world.

Questions about "Player 1." At the end of the study, participants were asked to type an open-ended response to the question, "What do you think about your interaction partner (Player 1)?" They were also asked whether or not they thought their partner thought they were real, and whether they believed their partner was real. Then they were told that their partner was not real, and were asked to report on a seven-point scale how surprised they were to find this out.

Procedure

All participants were recruited online on MTurk. A link was provided to the Qualtrics-based study. Participants were paid $1.50 for their time, with the added incentive of possibly winning a $5 bonus based on raffle tickets awarded during the economic deception game. Participants went through the economic deception game described in the pilot studies above. Next, they were presented with the picture and biography for all four characters, either one at a time, or arranged next to each other in a table-like format to facilitate direct comparisons. If the

characters were presented together, they were arranged on the screen left to right - administrator, doctor, quack, and teacher. When presented alone, they were also presented in this order, just one at a time.

The other measures outlined above were administered after the trait rating task. Each web page used contained a hidden timer that measured how many seconds were spent on that page.

Results

Participation Evaluation

In total, 307 people clicked the link to the Qualtrics survey for this study, 289 of which stayed in the study at least far enough to complete the economic deception game.

Participants were warned in the MTurk description not to use a smart phone or a tablet to participate in this study because their screens are too small to show all four characters side by side simultaneously, so participants whose metadata revealed that they had used one of these small screens to complete the study were excluded from analysis. Ninety-six percent (278) of eligible participants passed this criterion.

Participants were also quizzed on their retention of the content of the character descriptions in order to ascertain whether they carefully read these descriptions. They were asked two questions about the details of the characters' written descriptions, such which character had which hobby or which job. If they could not answer the two questions correctly, the participant was also excluded from analyses. Seventy percent (205) of the total eligible participants passed this criterion. This would be considered a high rate of exclusion for most university participant pool studies, but is not abnormal for a cognitively-demanding MTurk study – exclusion rates of 40% or more are not uncommon (Kaufman, Shulze, & Veit, 2011; Kittur, Chi, & Suh, 2008).

Ancillary Measures

Of the participants included in analyses, 22 (11.2%) went against Player 1's advice, demonstrating initial distrust.

One hundred twenty-six (64%) of participants reported using none of the pseudoscientific medical products in the survey, 53 (27%) reported using one, and 17 (9%) reported using more than one in the past year.

One hundred seventy-five of included participants (89%) reported using none of the pseudoscientific medical services in the survey, 18 (9%) used one service, and three (1%) used two in the past year.

Seventy-nine participants (57%) reported that they thought that Player 1 believed *they* were a real person, and 88 (45%) reported that they believed that *Player 1* was a real person. When it was revealed that Player 1 was actually a simulation, participants were fairly unsurprised ($M = 2.92$, $SD = 1.75$) on a scale of 1 (extremely unsurprised) to 6 (extremely surprised).

Character Trait Judgments

One hypothesis for this study was that all characters would receive more extreme ratings in the distrust condition when presented together than when presented apart, because with no other character available to find dissimilarities with, participants would instead contrast the character with his own group, most likely estimating him to be less stereotypical, as demonstrated in a past study by Posten and Mussweiler (2014). The extremity of trait judgments was calculated by taking the absolute value of the difference between each individual trait rating and the midpoint (3). Medically-relevant traits and non-medically-relevant traits were averaged separately to create two composite measures (one key dependent measure and one filler

measure). ANCOVAs were conducted on the two measures, for each of the four characters separately, in which effects of the partner and presentation conditions (alone vs. with others) were examined, controlling for reported use of pseudoscientific products and services. A secondary analysis looked at the direction of trait ratings, using a mixed factors ANOVA, comparing the doctor and quack only, with medically-relevant traits as the dependent variable.

Extremity of medically-relevant trait ratings. Contrary to expectations, the doctor received marginally more extreme trait ratings when presented *alone* (M = 1.80, SD = .32) than when presented *with others* (M = 1.70, SD = .43), $F(1, 190) = 3.46$, $p = .064$, $d = .28$. There was no significant main effect of partner, $F(1, 190) = 1.47$, $p = .227$, $d = .12$, nor was the predicted interaction significant, $F(1, 190) = 3.46$, $p = .064$, $\eta_p^2 = .01$. Use of pseudoscientific products and services did not have an effect, $F < 1$.

For the quack character, there was a significant interaction between partner and presentation on the extremity of the quack's trait ratings $F(1, 190) = 4.99$, $p = .027$, $\eta_p^2 = .026$, as predicted. However the direction of effects was also contrary to predictions. Trait ratings were more extreme in the *distrust* condition when the quack was presented *alone* (M = .93, SD = .57) than *with others* (M = .79, SD = .49). In the *trust* condition the direction of effects was normative – the direct comparison available in the *with others* condition appeared to make differences larger (M = .92, SD = .48) than if characters were presented one at a time, (M = .78, SD = .46) . However, decomposition of the interaction into simple main effects revealed that the simple effects of presentation were not reliable in either the trust or distrust conditions, $ps > .10$.

There were no significant effects of distrust or presentation on the extremity of the hospital administrator's trait ratings.

There were no significant main or interactive effects of partner or presentation on the extremity of the teacher's ratings.

Descriptive statistics for medically-relevant trait ratings can be found in Table 13.

Extremity of non-medically-relevant trait ratings. No effects were predicted on this composite measure, and indeed no significant main or interactive effects of partner or presentation were found for any of the four characters, *ps* > .20. However, use of CAM products was a significant covariate in several of the analyses, which was not true for the medical traits composite measure.

Use of CAM products was associated with more extreme generic ratings of the doctor, $F(1, 190) = 4.20$, $p = .042$, $\eta_p^2 = .02$, while use of CAM *services* was associated with *less* extreme ratings of the same character $F(1, 190) = 3.06$, $p = .082$, $\eta_p^2 = .02$. Examination of the average *directional* trait ratings indicated that the mean trait rating was slightly above the midpoint (3; $M = 3.50$, $SD = .42$). Thus more CAM products led to exaggeratedly positive impressions of the doctor on nonmedical traits, and more CAM services led to more muted impressions.

Similarly, use of CAM products led to more extreme ratings of the quack on medically unrelated traits, $F(1, 190) = 7.201$, $p = .008$, $\eta_p^2 = .037$. The quack was judged on average to be within one standard deviation of the midpoint on these traits ($M = 3.12$, $SD = .54$), and the use of CAM products correlated positively with trait ratings $r(195) = .27$, $p < .001$.

The same pattern held for the teacher. There were no effects of condition, but use of CAM products $F(1, 190) = 5.42$, $p = .021$, $\eta_p^2 = .03$ and services $F(1, 190) = 4.27$, $p = .022$, $\eta_p^2 = .02$ covaried significantly with the non-medically-relevant traits, with CAM product use associated with more extreme judgments, and the use of CAM services being associated with

less extreme judgments. The teacher was judged on average to be slightly more than a standard deviation from the midpoint for these traits (M = 3.48, SD = .45), so the greater extremity of ratings indicates more positive ratings on these traits.

Belief in partner was not a significant predictor of the extremity of ratings for any of the characters ($p > .05$).

Medical competence and skill. To examine directional effects, a 2(Partner: trustworthy vs. untrustworthy) x 2(Presentation: together vs. apart) ANCOVA with use of CAM products and services as covariates was performed on the average medically-relevant trait ratings for each of the characters. Ratings for the doctor were predicted to be lower in the distrust condition when the doctor was presented alone than when he was presented with the other characters, and ratings for the quack were predicted to show the opposite pattern – higher in the distrust condition when presented alone than when presented with others. Descriptive statistics for medically-relevant trait ratings of all characters can be found in Table 12.

For the doctor, there was a significant effect of presentation on medically-relevant trait ratings. Participants who saw him with the other characters rated him lower on these traits (M = 4.66, SD = .51) than those who saw him alone (M = 4.80, SD = .32), $F(1, 190) = 4.79$, $p = .030$, d = .34, contrary to expectations. There was no main effect of partner, $F(1,190) = .09$, $p = .770$, and the predicted interaction was not significant, $F(1,190) = 1.37$, $p = .243$. Use of CAM products and services were not significant covariates in this analysis.

For the quack, there were no effects of either partner or presentation on medically-relevant trait ratings. Contrary to prediction, the interaction was not significant either, $F(1,190) = 1.53$, $p = .217$. CAM product use significantly covaried with the traits $F(1, 190) = 16.51$, $p <$

.001, η_p^2 = .08, such that those who used more gave higher trait ratings. CAM service use was not a significant covariate.

Belief in partner was not a significant predictor of these ratings ($p > .05$).

Dispositional distrust and medical competence. MMV, along with trust in police, neighbors, relatives, and the government, was not found to correlate significantly with medically-relevant trait ratings for either doctor or quack (ps > .09).

Confidence in one's own judgments correlated positively with medically-relevant trait ratings of the doctor, $r(196) = .19$, $p = .007$, but not the quack, $r(196) = .045$, $p = .527$. Personal control correlated positively with medically-relevant trait ratings of both the doctor, $r(196) = .15$, $p = .037$ and the quack, $r(196) = .21$, $p = .004$. Self-trust also correlated positively with medically relevant traits of both the doctor, $r(196) = .22$, $p = .002$ and the quack, $r(196) = .18$, $p = .011$. As one might expect, trust in science correlated positively with medical trait ratings of the doctor, $r(195) = .20$, $p = .005$, but negatively with those of the quack, $r(195) = -.18$, $p = .011$. Trust in one's religion did not correlate with medical traits of the doctor, $r(196) = .12$, $p = .869$, but did correlate positively with medical traits of the quack, $r(196) = .32$, $p < .001$.

There were no specific hypotheses concerning the hospital administrator, but medically-relevant trait ratings followed an intuitive pattern. There was a marginal main effect of partner on the medically-relevant trait ratings of the hospital administrator, in which distrusting participants rated him slightly lower ($M = 3.87$, $SD = .61$) than trusting participants ($M = 3.95$, $SD = .56$), $F(1, 190) = 2.87$, $p = .092$, $d = .13$, and a marginal main effect of presentation, with ratings being lower when the administrator was presented with others ($M = 3.87$, $SD = .56$) than when he was presented alone ($M = 3.98$, $SD = .64$), $F(1, 190) = 2.75$, $p = .099$, $d = .19$, which is in line with expectations, since it invites comparison to an actual physician. There was a significant

interaction between these two variables, $F(1, 190) = 5.30$, $p = .022$, $\eta_p^2 = .03$, such that trusting participants rated him higher on the traits when he was presented alone ($M = 4.19$, $SD = .50$) than with others ($M = 3.85$, $SD = .56$) $F(1, 95) = 3.78$, $p = .010$, $d = .64$, and distrusting participants showed no simple effect of presentation. Use of CAM services and products did not significantly covary with trait ratings.

For the teacher, there was a significant main effect of presentation. The teacher was rated higher on medically-relevant traits when presented alone ($M = 3.65$, $SD = .49$) than when presented with others ($M = 3.49$, $SD = .51$), $F(1, 190) = 7.03$, $p = .009$, $d = .32$. The use of CAM products $F(1, 190) = 10.48$, $p = .001$, $\eta_p^2 = .05$ and CAM services $F(1, 190) = 7.18$, $p = .008$, $\eta_p^2 = .04$ were significant covariates in the analysis. CAM product use was associated with higher trait ratings, and CAM services were associated with lower trait ratings.

Direct Comparison of the Doctor Versus Quack Impressions. As noted above, no directional predictions were made about impressions of the administrator and teacher on trait ratings of medical competence and skill; the key predictions focused on the doctor and quack. A follow-up analysis directly comparing ratings of these characters was thus conducted: a 2 (Character: Doctor vs. Quack) x 2 (Partner: Trustworthy vs. Untrustworthy) x 2 (Presentation: Alone vs. Together) mixed factors ANOVA, with character as the within-subjects variable. As expected, there was a large main effect of character, with the doctor rated more highly on the traits ($M = 4.71$, $SD = .46$) than the quack ($M = 3.29$, $SD = .86$) $F(1, 192) = 481.94$, $p < .001$, $d = 2.06$.

There was a marginal interaction between character and presentation on impressions. Presenting the characters together with others resulted in lower ratings of the doctor (alone, $M =$

4.80, *SD* = .32; together, *M* = 4.68, *SD* = .51), and more positive impressions of the quack, (alone, *M* = 3.23, *SD* = .90; together, *M* = 3.33, *SD* = .83), $F(1, 192) = 3.25$, $p = .073$, $\eta_p^2 = .02$.

This two-way interaction primarily held for participants in the trust condition, a marginally significant 3-way interaction between character, partner, and presentation revealed, $F(1, 192) = 3.13$, $p = .078$, $\eta_p^2 = .02$. The convergence in trait ratings for the doctor and the quack when they were presented together was only apparent in the trustworthy partner condition as the means in Table M indicate.

Discussion

The predicted interaction between presentation and partner on the extremity of medically-relevant traits was not supported by the data for any of the four characters. Although it was statistically significant for the quack, the interaction effect was in the opposite direction than what was predicted. Likewise, the hypothesis that medically-relevant trait judgments of the doctor and the quack would converge among participants in the distrust condition who viewed the characters separately was also not supported by the data. In fact, trends in the data pointed to this convergence being somewhat more likely in the *trust* condition.

Taken at face value, these results represent a failure to find evidence for the effect of distrust on trait judgments relevant to medical competence, regardless of whether they are presented to participants together or one-at-a-time. As with Study 1, it could be that this particular manipulation was not strong enough to produce an increase in contrast sensitivity detectable by this dependent measure, or it could be that the basic cognitive effects observed by other researchers do not extend to character judgments as I theorized they might. And again, we must wonder whether "distrust" can be understood as one construct, or if different ways of promoting distrust are actually acting on adjacent, but still distinct, processes.

On CAM Use

Participants were asked to report what CAM products and services they had used in the past year so that the total number of products or services used could be used to control for favorability toward alternative medicine – which could have confounded judgments of the quack vs. the doctor. While CAM product use did correlate with judgments of the quack's medical competence, they did not correlate with the same judgments pertaining to the doctor. This may suggest that CAM use corresponds with a general openness to alternative treatments that does not necessarily include a devaluing of science-based medical practice.

CAM service use did not covary with judgments of medical competence for either character, but so few participants reported the use of CAM services that it would not be appropriate to draw any strong conclusions about this particular relationship based on the data at hand.

So far, manipulated distrust has not been shown to have the predicted influence on either the judged likelihood of statements, or the judged competence of persons. The next question I will ask is whether manipulated distrust can influence the favorability or unfavorability of cognitive responses to persuasive arguments.

Chapter 10

Experiment 3

This study will attempt to explain one of the more stereotypical qualities believed to be associated with distrust – resistance to attitude change – in terms of elevated contrast sensitivity.

So far we have found that experimentally manipulated distrust failed to make novel claims more believable, and also did not appear to promote individual-group contrasts in stereotypical judgments. It is possible then, that distrust may simply make people more resistant to attitude change.

According to the Elaboration Likelihood Model (Petty & Cacioppo 1986), when the issue at hand is personally relevant, people become more likely to evaluate arguments more carefully, generating cognitive responses, and rely less on heuristics. This leads them to favor strong arguments over weak arguments. Theoretically, this is because strong arguments generate more favorable and fewer unfavorable thoughts than do weak arguments. However, if a state of distrust increases the availability of contrasting information (Schul et al., 2004), then involved message recipients may generate more contrarian cognitive responses, resulting in a less favorable attitude toward even relatively strong arguments.

This study makes use of a topic that college students tend to have strong feelings about - comprehensive senior examinations as a requirement for graduation (Petty & Cacioppo, 1979). Students are generally understood to be biased against such a requirement, though they can become more favorable toward the idea if they are given strong arguments to justify it. In the current study, participants in the distrust condition were predicted to resist attitude change and report more negative responses to the pro-exam arguments, to report more positive responses and fewer negative responses to an anti-exam argument, and to report more cognitive responses,

overall. An anti-exam argument was added so that it would be possible to measure participants' favorability toward an opposing argument. Presumably, if distrust facilitates the activation of contrasting information, then a statement that contrasts with the previously-read statements might be viewed more favorably by distrusting than trusting participants.

Method

Participants and Design

One hundred sixty-nine participants clicked the survey link, and 161 stayed in the study long enough to complete the critical measures. Participants were 30% male, 70% female, 70% White, 5% Black, 17% Asian, and 3% wrote-in that they identified as American, human, Indian, or Native American. Despite the elimination of eight participants, they were evenly distributed between conditions – 49.7% were assigned to a trustworthy simulated "partner," and 50.3% to an untrustworthy one.

Materials and Procedure

Acquisition and adaptation of materials. Six strong arguments in favor of implementing senior comprehensive examinations gleaned from classic elaboration likelihood studies were presented to students (See arguments listed in Appendix W; adapted from Petty & Cacioppo, 1979), and a seventh statement was written in opposition to senior comprehensive exams.

Distrust manipulation. The economic deception game detailed in the pilot studies, and employed in Study 2 was used again in this study. Participants were randomly assigned to condition by the computer. In both conditions, a description of the procedure was provided. Participants were then quizzed for comprehension of that description. If they picked the wrong

answer, they were given another chance to read the description and answer the question again before moving on.

Presentation of and response to arguments. After the economic deception game, participants were presented with all seven arguments, one at a time (the anti-exam argument was presented last). First, the argument was presented, with instructions telling participants to simply read it. On the next page, the argument was presented again, and participants were asked to "list ANY thoughts that occurred to you while you were reading this statement." They were given a maximum of 120 seconds to give their response before the page automatically moved them to the next page. Participants were asked to separate their individual thoughts by starting a new line.

On the next page after that, participants were again presented with the same argument and asked to rate their agreement with the statement on a seven-point scale. This entire process was repeated for each of the seven arguments.

Other measures. After these critical measures, other measures of general trust and openness were administered, as in the previous two studies. Confidence and Control, the Revised Health Care System Distrust Scale (Shea et al, 2008), measures of trust in neighbors, relatives, self, science, police, and religion, and government, the IPIP scale of intellectual openness, and demographic measures were included.

Questions about "Player 1." Participants completed the same questions about their thoughts about and belief in "Player 1" as in the previous studies – their opinion of Player 1, whether Player 1 believed they were real, whether they believed Player 1 was real, and their level of surprise that Player 1 was not real.

At the end, participants were given the debriefing materials explaining the deception used in the study and the hypotheses of the study, and were thanked for their participation.

Results

All 161 participants who stayed in the study long enough to take the critical measures were included for analyses. Examining the initial distrust responses revealed low initial distrust. Of the participants included in analysis, six (3.8%) went against Player 1's advice, indicating an initial tendency toward distrust.

Cognitive responses. Consistent with previous studies, two coders coded each thought as either favorable, unfavorable, or neutral. Interrater reliability was generally high (see Appendix W for Kappa scores). Analyses were conducted on composite measures averaging how many favorable, unfavorable, neutral, and total thoughts were calculated in response to the six pro-exam arguments. Separate analyses were conducted on the number of thoughts in each of these categories generated in response to the final, and only, anti-exam argument.

Separate t-tests (See Table 15) were performed on each of the thought measures in order to test the hypothesis that participants in the untrustworthy partner condition would report more unfavorable thoughts, and more thoughts total. There were no significant effects of partner on the mean positive, negative, neutral, or total responses per pro-exam argument, The same was true for responses to the one anti-exam argument (absolute value of all t-tests was <1.5, all *p*s > .1).

For exploratory purposes, responses to individual arguments were analyzed. Consistent with predictions, participants with a trustworthy partner reported more positive thoughts (M = .67, SD = .94) in response to the sixth argument (this argument suggested that comprehensive exams could reduce tuition rates by increasing alumni support) than did participants with an untrustworthy partner (M = .41, SD = .65) $t(138.27) = 2.01$, $p = .047$, $d = .32$ (equal variances not assumed). Descriptive statistics for all arguments can be found in Table 14.

Additionally, an exploratory 2 (position: pro-exam vs. anti-exam) x 2 (partner: trustworthy vs. untrustworthy) x2 (Believe vs. Disbelieve in partner) mixed ANOVAs were performed with MMV as a covariate representing dispositional distrust for each category of response (favorable, unfavorable, neutral, and total thoughts, and agreement with the arguments), and revealed one three-way interaction between position, partner, and belief on total cognitive responses, $F(1, 152) = 3.89$, $p = .05$, $\eta^2 = .025$ in which participants who believed in their partner had fewer cognitive responses to the anti-exam argument if their partner was untrustworthy, and had more cognitive responses to the anti-exam argument when their partner was trustworthy (See Figure 2), and participants who believed in their partner generally agreed more ($M = 4.49$, $SD = .83$) than those who did not ($M = 4.05$, $SD = .97$) with all statements, regardless of condition, $F(1, 152) = 3.90$, $p = .05$, $d = .49$. No other main or interaction effects were observed. MMV did not covary in any of these analyses.

Agreement with the arguments. Agreement with pro-exam arguments was calculated by averaging the reported agreement with each pro-exam argument. There was no significant effect of partner on mean agreement with the pro-exam arguments, $t(157) = -.582$, $p = .561$, or agreement with the anti-exam argument $t(155) = .65$, $p = .517$.

Having a trustworthy partner was also marginally associated with reporting greater agreement with the sixth argument ($M = 4.41$, $SD = 1.46$) than having an untrustworthy partner ($M = 3.96$, $SD = 1.69$), $t(154) = 1.78$, $p = .078$, $d = .29$.

Effects of Initial Distrust

For exploratory purposes, t-tests on all measures were also performed to look for the effects of initial distrust. As was the case with the partner effect analyses above, the only significant difference occurred with the sixth argument. Contrary to my theoretical rationale,

participants who were initially distrusting had fewer negative thoughts in response to argument six (M = .17, SD = .41) than those who were initially trusting (M = .91, SD = 1.14) t(8.48) = 3.90, p = .004, d = .87 (equal variances not assumed).

There were also marginal effects of initial distrust on positive thoughts in response to the fifth argument – initially trusting participants had more positive thoughts (M = .97, SD = 1.04) in response to the argument than did initially distrusting participants (M = .17, SD = .41) t(157) = 1.87, p = .063, d = 1.01 – and agreement with the sixth argument – initially trusting participants agreed less strongly with the argument (M = 4.14, SD = 1.56) than initially distrusting participants (M = 5.33, SD = 2.07) t(154) = -1.82, p = .071, d = -.65.

Note that due to the small number of initially distrusting participants in this sample (6), these findings should be interpreted with caution.

Dispositional Distrust Effects

Medical mistrust. An exploratory correlative analysis found that neither subscale of medical mistrust correlated with responses to the arguments (all absolute rs <.13).

Other measures of trust. Self-reported confidence correlated positively with positive responses to pro-exam arguments r(157) = .19, p = .019. Trust in relatives was positively correlated with positive responses to pro-exam arguments, r(157) = .16, p = .048, and negatively correlated with neutral responses to pro-exam arguments, r(157) = -.16, p = .047. Trust in religion was negatively related to negative responses to pro-exam arguments, r(157) = -.16, p = .046.

Intellectual openness. IPIP intellectual openness was positively correlated with negative responses to pro-exam arguments, r(157) = .20, p = .013, and positively correlated with total responses to pro-exam arguments, r(157) = .16, p = .044.

Discussion

The predicted pattern of results – that participants paired with untrustworthy partners would produce more negative responses to pro-exam arguments, more positive responses to an anti-exam argument, and provide overall more cognitive responses – was not found in the results of this study. That said, participants who had an untrustworthy partner did report significantly fewer positive thoughts in response to the sixth argument, which stated that instituting comprehensive exams would increase alumni donations, and thus, lower tuition. They also reported less agreement with this argument than did participants with a trustworthy partner. So, while distrust did not promote a bias against most of the arguments, it did seem to bias participants against this particular argument.

Argument six might have been particularly susceptible to the economic deception game's manipulation of trust because the manipulation specifically did have to do with a monetary prize, and simulated an interaction with another student. Being so recently deceived by a fellow student in a game for a chance at winning money might make student participants particularly more cynical about the prospect of alumni donating enough for tuition costs to go down.

By the same token, it is also possible that, upon having a trustworthy fellow student try to help them get a chance at winning the prize money, participants in the trust condition may have become more confident in the generosity of their school's alumni. This second explanation, though less in line with the theoretical idea that trust is normal and distrust a departure from the norm, is more supported by the data here, because participants in the distrust condition did not produce more unfavorable cognitive responses to the argument, only fewer favorable ones. This seems to suggest that participants in the trust condition were simply coming up with more positive responses to the argument than their distrusting counterparts.

At this point, manipulated distrust has not been shown to increase contrast sensitivity, in the realms of the judged likelihood of truth statements, the judged competence of professionals, or in the generation of cognitive responses to counter-attitudinal arguments. The next set of studies will address the other hypothesized effect of distrust on cognition - cognitive flexibility.

Chapter 11

Experiment 4

The three studies above have addressed contrast sensitivity as a mechanism through which distrust could change the way persuasive messages are interpreted. The main thrust of that hypothesis was that if a state of distrust facilitates the search for contrasting information or for contradictions to the present information, then by way of the availability heuristic (Tversky & Kahneman, 1973) – which states that concepts which are easier to bring to mind are assumed to be more likely to be true, or more likely to occur – opposing messages should seem more convincing than they otherwise would.

That hypothesis did not find strong support in those studies, though there were some interesting findings concerning dispositional distrust of the medical establishment.

The present study, as well as study 5, will address the hypothesized cognitive flexibility associated with a state of distrust. If, as some studies seem to suggest (Mayer & Mussweiler, 2011; Schul, Mayo, & Burnstein, 2008), distrust promotes cognitive flexibility, then that may mean that people in a state of distrust can be expected to be more oriented toward alternatives to the norm, or may be more likely to believe that information is being concealed from them.

In this study, cognitive flexibility was operationalized as the number of unique completions of two-letter word stems. Distrust was induced via the imposter detection task. Participants in the impostor-detection condition were expected to generate more complete words per stem than participants who were asked to judge each respondent's spontaneousness.

Method

Participants and Design

Participants were recruited through MTurk. They were offered $1.50 for their participation if they completed the study and entered the confirmation code given to them at the end. Two hundred eighty-five people clicked the survey link, 274 stayed in the study long enough to complete the manipulation, and 242 passed the attention check item that asked them to "Please answer "Disagree" for this item to confirm you are reading these questions." These 242 participants were included in the analyses.

The sample was 43% male, 56% female, and .4% transgender. Eighty percent of participants identified as White, 12% Black, 6% Asian, .4% Native Hawaiian/Pacific Islander, 5% Latino, and 2% wrote in that they were Indian, mixed, or Native American. Participants were 36 years old on average.

Participants were randomly assigned to either the impostor detection or spontaneity rating condition.

Materials and Procedure

Upon beginning the study, participants were presented with an information sheet summarizing the things that they would be asked to do in the study.

Imposter detection task. This experiment used the same athlete impostor detection judgment task and control judgment task described in the pilot studies above, and in study 1. In the pilot studies, this task was the most effective for MTurk workers, despite the fact that it directly referenced University of Connecticut athletics. Participants were distributed evenly between conditions – 51% (123) were assigned to the impostor detection condition, and 49% (119) were assigned to the spontaneity judgment condition.

Word Stems. After either completing the imposter detection task, or the spontaneity-rating task, participants were introduced to the word stems they would be completing. Each stem consisted of two letters, and participants were given 30 seconds to type as many words that began with those two letters as they could. At the end of the 30 seconds, the next word stem was presented. Word stems were generated for this study by me. Participants were given two practice word stems before starting the proper set of 25 word stems (See Appendix X for a complete depiction of the presentation of word stems).

Other trust measures. As in the above studies, other measures of trust were included, including the confidence, control, Revised Health Care System Distrust Scale (Shea et al, 2008), and measures of trust in neighbors, relatives, self, science, police, and religion, and government. These dispositional trust measures were used for the purposes of correlational analysis of the number of completed words.

At the end of the study, participants were debriefed – first they were asked what they thought the study was about (none guessed correctly), were told about the deception used in the study and hypotheses, and were asked if they had had any problems in completing the study.

Results

Word stems were coded by counting the responses to each word stem for each participant. A response was considered to count if it was a real word, started with the appropriate letters, and was unique to the other responses for that participant and word stem. Occasionally, a participant would respond by typing only the letters that followed the word stem (e.g. for the stem: El____, they would respond, “ephant”). These responses were counted as if they contained the word stem.

A t-test was used to test for the effect of the judgment type (experimental condition) on the average number of words per stem. There was no effect of judgment type on the average number of words generated on the test trials, $t(240) = .64$, $p = .526$. (Individual trial responses are presented in Table 16.)

For exploratory purposes, the number of words generated on the two practice trials were also analyzed. On the very first practice word stem, participants who were asked to identify impostors generated fewer responses ($M = 4.70$, $SD = 2.65$) than participants who were asked to make spontaneity judgments ($M = 5.40$, $SD = 3.19$) $t(240) = 1.87$, $p = .063$, $d = .24$ contrary to predictions.

Effects of Dispositional Distrust on Word Generation

Correlational analyses were conducted in order to look for associations between dispositional indicators of distrust and the generation of word stem responses.

Neither of the factors of the Revised Health Care System Distrust Scale (Shea et al, 2008) correlated with total stems completed. MMV: $r(242) = .01$; MMC: $r(242) = .001$, both $ps > .9$. There were some positive, weak correlations with other measures of trust – trust in neighbors $r(242) = .14$, $p = .028$, relatives $r(241) = .21$, $p = .001$, and science $r(242) = .13$, $p = .04$, all involved increased total number of responses to word stems being associated with higher levels of trust. This finding was also in the opposite of the theorized direction.

Discussion

The main predicted effect was not found. There was no difference between experimental conditions in the average number of responses per word stem. Responses to the very *first* word stem did show an effect of the distrust manipulation, albeit not in the predicted direction – the control group produced more solutions than did the distrust group.

Dispositional measures of distrust generally did not correlate with the number of responses, or correlated weakly, again with greater trust being associated with more responses, not less.

Once again, my predictions were not supported by the data. While responses to word stems represent creativity in the sense that each response is both novel and correct (Sternberg & Lubart, 1999), it is also a task that does not require much in terms of judgment and reasoning. The next experiment tackles the way people evaluate and reconcile conflicting pieces of information under distrust.

Chapter 12

Experiment 5

Distrust is often thought of as being a generally contrarian or conservative mindset, but my interpretation of the literature also implies that it may also prompt people to become more creative in the way they combine information.

Pseudoscientific and conspiracy-oriented beliefs are often associated with distrust. Goertzel (1994) found a lack of interpersonal trust was one of the reasons that members of racial minorities are more likely to believe in conspiracy theories, and a common thread in anti-vaccine (Wolfe et al., 2002) and climate change (Lewandowski, 2013) literature is that the scientific establishment cannot be trusted. At the same time, experimental studies on distrust have shown that having a distrusting mindset can make a person more sensitive to information that conflicts with what they see (Schul et al., 2004), and more sensitive to contrasts (Posten & Mussweiler, 2013). Distrust has also been shown to be associated with increased cognitive flexibility, allowing for greater creativity (Meyer & Mussweiler, 2011), and enhanced performance on puzzles that require participants to think in unconventional ways (Schul et al., 2008).

This increase in creativity and the salience of contrasts may allow for the possibility that when someone is distrusting, they are able to maintain a position (e.g., opposition to vaccines, or belief that there is no climate change) by seeing pieces of information that oppose their position as non-representative exceptions, rather than disconfirming evidence to their position. In addition, a boost in cognitive flexibility may allow them to find creative ways to "bend" the facts into their favor by imagining possible conditions under which the encountered information does not contradict their position.

This may be a counter-intuitive interpretation of the contrast sensitizing effects of distrust. After all, if distrust increased sensitivity to contrasts, that could mean that they would be

more likely to see the inconsistencies between their beliefs and the reality they witness. However, distrust is not a value-free state. It is fundamentally a way of thinking that is more alert to attempts at manipulation or deception against the self. Someone who distrusts their environment is not necessarily distrusting of their own feelings and beliefs. So their distrust, and the mechanisms that come with it are more likely to be directed outward than directed inward.

Furthermore, judging two obviously conflicting statements to be in conflict with each other would be a routine, non-creative, compliant way to approach the problem. In a situation such as this, a distrusting person is predicted to find an alternative answer that contrasts with the surface-level implications of the information presented to them.

The present study will address the simple idea that distrust can make two conflicting statements seem more compatible than they are. Participants will be given three different kinds of statement pairs – statements that are unrelated to each other, statements that conflict without directly contradicting each other, and statements that directly contradict one another. Participants will be asked to rate how compatible each pair of statements are with each other, and then will be asked to describe what conditions would be necessary to make them compatible. Beforehand, participants will either be induced to distrust, or will be given a filler task. It is predicted that distrusting participants will judge the unrelated and directly contradicting statements similarly to control participants, but that moderately conflicting statements will be more likely to be judged to be compatible by distrusting participants than by control participants, and that explanations for their compatibility will be more complex.

Method

Participants and Design

Participants were recruited through the University of Connecticut participant pool. They were given a survey link to follow to participate in the experiment. Two hundred eleven students clicked the survey link, with 12 of them dropping out before the manipulation was completed. The remaining 199 were included in analyses to the degree that they provided data post-manipulation.

This sample was 28% male, 69% female, and 1% transgender. Sixty-nine percent identified themselves as White, 8.5% as Black, 15.5% Asian, .5% Native Hawaiian/Pacific Islander, 6.5% Hispanic/Latino/a, and 2.5% identified themselves as American, Arab, Asian American, Indian, or Native American. They were 19 years old on average.

Participants were assigned randomly to either the trustworthy partner or untrustworthy partner conditions, and each participant responded to all three statement pair groups – unrelated pairs, conflicting pairs, and directly contradicting pairs.

Materials and Procedure

Economic deception game. The economic deception game detailed in the pilot studies, and employed in Study 2 and 3 was used again in this study. Participants were randomly assigned to condition by the computer. In both conditions, a description of the procedure was provided. Participants were then quizzed for comprehension of that description. If they picked the wrong answer, they were given another chance to read the description and answer the question again before moving on.

Pairs of statements. After the between-subjects manipulation, a list of 6 statement pairs was presented to participants. Two of these pairs consisted of two statements that do not conflict

with each other at all and are unrelated to each other, two pairs consisted of statements that conflict, but do not directly contradict each other, and a third set of two pairs consisted of statements that directly contradict each other (See Appendix Y).

Each pair of statements was followed by the question, "Regardless of whether you believe it, assume for a moment that the first statement is true. Given that, how possible or impossible is it that the second statement is also true?" Participants responded to each pair on a scale of *completely impossible* (-10) to *completely possible* (10). This question was followed by the question, "Regardless of how you answered the previous question, please explain in your own words what conditions would be necessary to make these statements completely compatible with each other. If the statements are entirely compatible already, please say so."

Other trust measures. As in the above studies, other measures of trust were included for exploratory purposes, including the confidence, control, Revised Health Care System Distrust Scale (Shea et al, 2008), and measures of trust in neighbors, relatives, self, science, police, and religion, and government.

Questions about "Player 1." As in the past studies using the same distrust manipulation, participants were asked for their thoughts about "Player 1"– their opinion of Player 1, whether Player 1 believed they were real, whether they believed Player 1 was real, and their level of surprise that Player 1 was not real.

At the end, participants were given the debriefing materials explaining the deception used in the study and the hypotheses of the study, and were thanked for their participation.

Results

Economic Deception Game

Students were distributed evenly between conditions – 100 were assigned a trustworthy partner, and 99 were assigned an untrustworthy partner. Ninety percent of the sample demonstrated comprehension of the instructions on the first try, and 96% did so by the second try. Nine percent also chose to go against their partner's advice, suggesting initial distrust.

Thirty-eight percent of participants thought that "Player 1" believed that they, the participant, were real, and 27% reported believing "Player 1" was real. On a 6-point scale from extremely unsurprised (1) to extremely surprised (6), participants averaged 1.90 (SD = 1.27). Participants who believed that their partner was real (M = 3.28, SD = 1.39) were more surprised than those who did not (M = 1.38 SD = .71) $t(62.56) = 9.48$ (equal variances not assumed), $p < .001$, $d = 1.72$.

Coding Typed Responses

Open-ended responses to the statement pairs were copied into a document and word counts for each response were generated using word processing software. Number of sentences were counted manually. Composite measures for number of words and sentences and possibility judgments were created by separately averaging the two responses for the unrelated, conflicting, and contradicting pairs (See Table 17 for descriptive statistics).

Effects of Pairing and Partner

For each dependent measure, a 2 (partner: trustworthy vs. untrustworthy) x 3 (Pairing: unrelated, conflicting, contradicting) mixed ANOVA was performed, with pairing type as the within-subjects variable.

Word counts. There was no main effect of partner on the number of words used, $F(1, 197) = .13$, $p = .724$, $\eta_p^2 = .001$. There was a main effect of pairing type $F(2, 394) = 10.66$, $p < .001$, $\eta_p^2 = .05$. Follow-up analyses of differences between the pairing types using the Bonferroni adjustment revealed that conflicting pairs were reconciled using significantly more words ($M = 19.15$, $SD = 13.48$) than both unrelated pairs ($M = 15.79$, $SD = 10.94$), $p < .001$, $d = .27$ and directly contradicting pairs ($M = 16.70$, $SD = 11.76$), $p < .001$, $d = .19$. There was no significant interaction of partner and pairing on number of words generated $F(2, 196) = .22$, $p = .80$, $\eta_p^2 = .002$.

Sentence counts. There was no main effect of partner on the number of sentences used, $F(1, 197) = .09$, $p = .761$, $\eta_p^2 < .001$, nor was there a main effect of pairing $F(2, 394) = 1.79$, $p = .17$, $\eta_p^2 = .01$, nor was the partner x pairing interaction significant $F(2, 394) = 1.56$, $p = .21$, $\eta_p^2 = .01$.

Possibility judgments. For possibility judgments – for which participants were asked to assume the first statement was true, and rate the second statement's possibility of being true given that assumption – there was no main effect of partner $F(1, 194) = .99$, $p = .32$, $\eta_p^2 = .01$. There was a main effect of pairing $F(2, 388) = 106.32$, $p < .001$, $\eta_p^2 = .35$. Bonferroni contrasts revealed that unrelated statements were judged to be significantly more possible ($M = 3.88$, $SD = 5.46$) than conflicting ($M = -2.76$, $SD = 5.46$) $p < .001$, $d = 1.22$ or contradicting statements ($M = -3.60$, $SD = 5.82$) $p < .001$, $d = 1.33$. There was not a significant interaction of partner and pairing on possibility ratings, $F(2, 388) = .24$, $p = .77$, $\eta_p^2 = .001$. Results for words, sentences, and possibility judgments by pairing and partner type can be found in Table 18.

Including Belief in Player 1

In case the participants' belief in the conceit of the manipulation was an important factor, the same set of ANOVAs were performed, including reported belief that Player 1 was a real person. This amounted to a 2 (partner: trustworthy vs. untrustworthy) x 2(belief: yes vs. no) x 3 (Pairing: unrelated, conflicting, contradicting) mixed ANOVA with pairing type as the within-subjects variable. Six participants failed to respond to the question of whether they believed their partner was real, so the means and standard deviations have changed slightly from the previous analysis.

Word counts. Again, there was no main effect of partner on the number of words used, $F(1, 189) = 1.01$, $p = .317$, $\eta_p^2 = .005$, and there was a main effect of pairing type $F(2, 378) = 8.97$, $p < .001$, $\eta_p^2 = .05$. Follow-up analyses of differences between the pairing types using the Bonferroni adjustment revealed that conflicting pairs were reconciled using significantly more words ($M = 19.52$, $SD = 13.44$) than both unrelated pairs ($M = 15.79$, $SD = 10.77$), $p < .001$, $d = .27$ and directly contradicting pairs ($M = 16.70$, $SD = 11.76$), $p < .001$, $d = .19$. There was no significant interaction of partner and pairing on number of words generated $F(2, 189) = .01$, $p = .936$, $\eta_p^2 < .001$.

There was no main effect of belief in one's partner, $F(1, 189) = .01$, $p = .904$, $\eta_p^2 < .001$, but there was a significant interaction between partner type, and belief in that partner $F(1, 189) = 5.07$, $p = .026$, $\eta_p^2 = .03$. While the simple effects of partner type did not attain significance, the data trended in a manner consistent with predictions for effective manipulations, namely: when participants believed their partner was real, they used more words when that partner was untrustworthy ($M = 20.17$, $SD = 14.21$), than when their partner was trustworthy ($M = 14.70$, $SD = 7.38$) *ns*, $d = .48$. When they did *not* believe their partner was real, a weaker effect in the

opposite direction was found. Participants who interacted with an untrustworthy partner used fewer words (M = 16.59, SD = 9.31) than those who interacted with a trustworthy partner (M = 18.68, SD = 10.95) *ns*, d = .21.

Sentence counts. There were no main effects for sentence count, though there was a significant crossover interaction between partner and belief that the partner was real, $F(1, 189) = 9.42$, $p = .002$, $\eta_p^2 = .05$. Investigating the interaction further revealed a similar result to that of the word counts. When they believed their partner was real, participants in the distrust condition used more sentences (M = 1.40, SD = .53) than those in the control condition (M = 1.08, SD = .31). Though the difference did not reach statistical significance, the effect size of this difference was large, d = .73. Conversely, for participants who did *not* believe their partner was real, those in the distrust condition used slightly fewer sentences (M = 1.23, SD = .37) than did those in the control condition (M = 1.34, SD = .48). This difference was not significant, and the effect size was small d = .25 (see Figure 2).

Possibility judgments. For possibility judgments – for which participants were asked to assume the first statement was true, and rate the second statement's possibility of being true given that assumption – there was no main effect of partner $F(1, 189) = .14$, $p = .71$, $\eta_p^2 < .01$. There was a main effect of pairing $F(2, 378) = 81.95$, $p < .001$, $\eta_p^2 = .30$. Bonferroni contrasts revealed that unrelated statements were judged to be significantly more possible (M = 3.88, SD = 5.46) than conflicting (M = -2.76, SD = 5.46) $p < .001$, d = 1.22 or contradicting statements (M = -3.60, SD = 5.82) $p < .001$, d = 1.33. There was no significant main effect of belief in partner $F(1, 189) = .187$, $p = .67$, $\eta_p^2 < .01$, nor did belief interact with the trustworthiness of the partner $F(1, 189) = 1.73$, $p = .19$, $\eta_p^2 = .01$.

Dispositional Distrust

Some significant correlations were found between some of the dispositional or transient trust measures and the dependent measures in exploratory analyses. Individuals who had lower feelings of personal control generated somewhat more words $r(194) = -.16$, $p = .029$ and sentences $r(194) = -21$, $p = .003$ on conflicting statements than those with higher ratings of control. Lower feelings of control was similarly associated with generating more sentences to resolve contradicting statements, $r(194) = -.24$, $p = .001$.

Having more judgmental uncertainty was associated with judging contradicting statements as more compatible, $r(194) = -.17$, $p = .020$.

Trust in science correlated positively with both words $r(194) = .23$, $p = .001$ and sentences $r(194) = .148$, $p = .038$ used in response to conflicting statements, and negatively with possibility judgements of contradicting statements $r(194) = -.29$, $p < .001$.

Additionally, those who were politically liberal used more words to resolve conflicting statements than those who were politically conservative, $r(194) = -.15$, $p = .043$.

No other measures were significantly correlated with the dependent measures.

Discussion

The prediction that moderately conflicting pairs of statements would be uniquely viewed as more compatible compared to other nonconflicting or directly contradicting pairs of statements under distrust did not entirely hold, but the minority of participants in the distrust condition who believed their partner was a real person did use both more words and more sentences generally in their efforts to reconcile contradictory statements than control participants, suggesting that they did engage in more cognitive elaboration and, possibly, creativity in their

responses. Believers did not, however judge paired statements to be any more compatible than did control participants or disbelieving participants.

This suggests that while an experience with an untrustworthy partner may make people apply more thought in evaluating the compatibility of two statements, they ultimately arrive at the same conclusion as those who have a trust-affirming experience.

The correlations found between the main dependent variables and the measures of personal control, judgmental uncertainty, trust in science, and political leanings seem to have more to do with personality traits that may be adjacent to the concept of trust, but fairly independent of it, giver that the general trust/distrust measures did not correlate with them at all.

Having an interaction with a person in which they behave in an untrustworthy way toward you may make you evaluate information more carefully, but does not, at least in the context posed by this experiment, mean that you will come to a dramatically different conclusion about that information. It is possible, given the evidence found here for greater cognitive elaboration under distrust, that a difference in compatibility judgments may have been observed if participants' cognitive resources had been limited, forcing them to rely on peripheral processes to make their judgments. The findings in this study appear to support a skeptical response being triggered by distrust, but not one suggestive of a low-grade form of paranoid thinking.

Chapter 13

Experiment 6

The apparent association of a diffuse state of distrust with an increased sensitivity to contrast (Schul et al., 2004; Posten & Mussweiler, 2013) and increased cognitive flexibility (Meyer & Mussweiler, 2011; Schul et al., 2008) imply that when people are distrusting, they are more attuned toward detecting evidence that things are not quite what they seem, and that they are on the lookout for alternative interpretations of the information available to them, which could reveal possible ulterior motives, undesirable outcomes, and alternative courses of action.

This shift in orientation can effectively make the distrusting person more critical of the messages presented to them, but at the same time, according to my theory, it would seem to systematically bias the person in favor of counterarguments that cite contradictory anecdotes, conflicts of interest, side effects, and alternative routes to the desired outcome. These features seem to be on full display in anti-vaccination propaganda. Websites opposed to vaccinations often contain statements that diseases have declined independently from the rise in vaccination, criticisms of the profit motives of doctors and pharmaceutical companies, anecdotes about illnesses of all kinds occurring soon after vaccination, and alternative methods for avoiding sickness such as hand washing and nutrition (Wolfe et al., 2002).

The present study makes use of a more benign topic – tire rotation – to demonstrate this model of distrust's effects on message processing. Participants were either primed to distrust or not, and were presented with two essays in a point-counterpoint format. The first essay explains the benefits of regular tire rotation, and explains that it is recommended by the National Institute for Automotive Service Excellence (ASE). The second essay argues that ASE has a vested interest in having people spend more money on auto maintenance (ulterior motive), that the

benefits of tire rotation don't justify the cost in time, money, or personal risk (side effects), and provides advice on how to drive in a way that wears tires evenly (alternatives).

Participants were asked how much they agree with each of the essays, then asked specifically how much they agree with each point made by each essay. It was predicted that distrusting participants would be more receptive to the counter-arguing essay than control participants, and will also find the individual counter-arguments more convincing.

Method

Participants and Design

Participants were University of Connecticut undergraduates recruited through the University participant pool. Upon signing up for the experiment, they were provided with a link to the qualtrics survey. Seventy-eight students clicked on the survey link, and of those, 65 stayed in the study long enough to complete the distrust manipulation, which was the first step of the study.

The sample for this study was 26% male, 72% female, and 2% transgender. They were 69% White, 8% Black, 15% Asian, 2% Native Hawaiian/Pacific Islander, and 11% Hispanic/Latino/a. They were 19 years old on average.

The participants were assigned randomly to interact with either a trustworthy or untrustworthy simulated partner in the economic deception game, before reading and answering questions about two point-counterpoint essays concerning the pros and cons of tire rotation.

Materials and Procedure

Economic deception game. The same economic deception game used in previous studies was used in this one. Sixty-three of the 65 participants passed the comprehension quiz on the instructions on the first try, and the other two passed on the second try.

Pro-tire rotation passage. Next, participants were given a short pro-tire rotation message to read. The message was designed to emphasize the safety value and economic value of tire rotation, and to show that an authoritative professional organization supports it. Stylistically, it was intended to mimic the rhetorical style of typical public service messages:

> Tire rotation is an essential automobile maintenance practice. It ensures that your tires do not wear unevenly, which could reduce traction with the ground and create an unsafe situation on the road. Regular tire rotation has been shown to extend the life of tires and to help prevent cars from spinning-out on the freeway. The National Institute for Automotive Service Excellence (ASE) recommends that the typical driver rotate their tires with every other oil change.

Participants were then asked to rate their agreement with this message on a seven-point scale from "strongly agree" to "strongly disagree."

Anti-tire rotation passage. Immediately following the pro-tire message, participants viewed an opposing message, specifically to mimic the rhetorical style of vaccination denialists - emphasizing the profit motive of the ASE, improbable but dramatic risks, and the importance of common sense, while implying that good, responsible people don't need to take this precaution:

> You can't believe everything you read. Tire rotation might be helpful some of the time for some people, but most of the time, it is a costly, unnecessary inconvenience. It is important to remember the the ASE is a business that trains and certifies auto mechanics, and so it has a vested financial interest in more people rotating their tires more often. If

> you do decide to forego the financial strain of paying a mechanic to rotate your tires and try to do it yourself, you put yourself at risk of physical injury. The car could fall off of its supports, crushing any part of your body that happens to be underneath it – a high price to pay for unnecessary maintenance. If you are truly concerned about your tires wearing unevenly, just follow your common sense – don't accelerate too fast, don't slam on your brakes, and always return home the same way you came so that you balance the number of right and left turns. If you are driving sensibly and responsibly, there is no need to rotate your tires.

After reading this statement, participants were asked to report their agreement with the overall position of this second essay, and their agreement with each major point therein – that the ASE promotes tire rotation out of a desire for profit, that the monetary cost of professional tire rotation is not justified by its benefits, that the personal risk of do-it-yourself tire rotation is not justified by its benefits, and that responsible driving practices reduce the need for regular tire rotation.

Other trust measures. As in the preceding studies, other transient measures of trust were included just after the measures of agreement with the different points of the essay, including the confidence, control, measures of trust in neighbors, relatives, self, science, police, and religion, and government, and most importantly, the more stable Revised Health Care System Distrust Scale (Shea et al, 2008).

Questions about "Player 1." As in the preceding studies using the same distrust manipulation, participants were asked for their thoughts about "Player 1"– their opinion of

Player 1, whether Player 1 believed they were real, whether they believed Player 1 was real, and their level of surprise that Player 1 was not real.

At the end, participants were given the debriefing materials explaining the deception used in the study and the hypotheses of the study, and were thanked for their participation.

Results

Participants were included in the study if they passed the comprehension quiz for the economic deception game's instructions on the first or second try. Sixty-five participants met this criterion.

Economic Deception Game

Distribution between conditions. Participants were fairly evenly distributed between conditions, with 54% having a trustworthy partner, and 46% having an untrustworthy partner. Only 3% of participants showed initial distrust and went against Player 1's advice.

Belief questions. Due to experimenter error, 15 participants did not receive the questions about Player 1. When asked if they believed that Player 1 thought that they, the participant were real, 44% of those who received the question said yes. When asked if they believed that Player 1 was real, 28% said yes. On a six-point scale from extremely unsurprised to extremely surprised, participants reported being fairly unsurprised (M = 2.22, SD = 1.42) to discover their partner was not a real person.

Partner Effects

Agreement with pro-rotation message. An independent-samples t-test revealed no significant effect of partner on reported agreement with the first, pro-tire rotation message, $t(63) = .97$, $p = .34$, $d = .24$.

Agreement with anti-rotation message. Independent-samples t-tests were also performed for participants' agreement with the anti-tire rotation message.

Contrary to predictions, overall agreement with the message was found to be significantly lower among those paired with an untrustworthy partner (M = 3.23, SD = 1.55) than with a trustworthy partner (M = 4.14, SD = 1.33) $t(63) = 2.55$, $p = .013$, $d = .63$.

Agreement with the first point of the message – that the ASE's advice to regularly rotate one's tires was motivated by profit – was marginally lower for those with an untrustworthy partner (M = 4.13, SD = 1.46) than with a trustworthy partner (M = 4.80, SD = 1.59) $t(63) = 1.75$, $p = .084$, $d = .44$.

Agreement with the fourth point – that responsible driving practices reduce the need for tire rotation – was also significantly lower for those paired with an untrustworthy partner (M = 4.17, SD = 1.51) than with a trustworthy partner (M = 5.11, SD = 1.41) $t(63) = 2.61$, $p = .011$, $d = .64$.

The second point – that the monetary cost of tire rotation was not worth the benefits, did not show an effect of condition $t(63) = -.18$, $p = .86$, $d = .04$, nor did the third point – that the personal risks of tire rotation were not worth the benefits $t(63) = .58$, $p = .56$, $d = .15$.

Dispositional Distrust

Correlations with agreement with messages. As in the preceding studies, values-based medical mistrust again correlated with more of the dependent measures in this study, correlating positively with general agreement with the anti-rotation message argument $r = .36$, $n = 65$, $p = .004$, and with the first, profit-oriented argument $r(65) = .29$, $p = .020$.

Agreement with the first, pro-rotation message was positively correlated with confidence $r(65) = .25$, $p = .041$, and trust in the police $r(65) = .27$, $p = .029$.

Interaction Between Dispositional Distrust and Manipulated Distrust

That the effects of manipulated distrust and dispositional distrust had opposite effects on agreement with anti-rotation arguments prompted an exploratory multiple regression analysis to investigate the way these variables interact to affect persuasion.

General agreement with anti-rotation message. First, a linear regression was calculated to predict agreement with the message based on partner and MMV. The "partner" variable was dummy coded, with a 1 indicating an untrustworthy partner and a 0 a trustworthy partner. MMV was mean-centered before inclusion in the model (M = 3.01, SD = .64). A significant regression equation was found $F(2, 62) = 7.20$, $p = .002$, with an R^2 of .19. Predicted agreement was equal to 4.07 - .75(Untrustworthy Partner) +.73(MMV). Agreement was lower for participants paired with an untrustworthy partner by .75 units, and was .73 units higher (or lower) for every unit of MMV above the mean (below the mean). Both partner and MMV were significant predictors of agreement with the anti-rotation message.

The equation was then re-calculated, this time including an interaction term created by multiplying the dummy-coded partner variable by MMV. This second regression equation was also significant $F(3, 61) = 7.02$, $p < .001$ with an R^2 of .26. Predicted agreement with the message was equal to 4.11 – .67(Untrustworthy Partner) +.364(MMV) + 1.41(U.P. x MMV). Agreement was .67 units lower for participants paired with an untrustworthy partner, .364 units higher per unit of MMV, and an additional 1.41 units higher (or lower) for each unit above the mean (below the mean) of MMV for participants paired with an untrustworthy partner only. Partner was a marginally-significant predictor, the interaction term was a significant predictor, and MMV was not (See Figure 3 for a visual representation, and Table 19 for a table of the models of each argument in this study).

Agreement with pro-rotation message. The same regression analysis applied to participants' agreement with the pro-rotation message did not reveal main effects of partner, MMV, or their interaction on agreement with the pro-rotation message.

Individual counter-arguments. The same analysis strategy was used for the individual anti-rotation arguments.

Desire for profit. The model including both partner type and MMV was significant $F(2, 62) = 3.90$, $p = .025$, with an R^2 of .11. Predicted agreement with the argument that the ASE recommended tire rotation out of a desire for profit was equal to 4.74 - .53(Untrustworthy Partner) +.63(MMV). Agreement was lower for participants paired with an untrustworthy partner by .53 units, and was .63 units higher for each unit of MMV. MMV was a significant predictor, partner type was not. Adding the interaction term did not significantly improve model fit.

Unjustified financial cost and personal risk. Neither partner type, MMV, nor the interaction between them significantly predicted agreement with either of these arguments.

Personal responsibility. The model including only partner type as a predictor was significant $F(1, 63) = 6.84$, $p = .011$, with an R^2 of .10. Predicted agreement with the argument that responsible driving reduces the need for tire rotation was equal to 5.11 - .95(Untrustworthy partner). Agreement was .95 units lower for participants paired with an untrustworthy partner. Neither MMV nor the interaction between partner type and MMV significantly improved the model when added.

Belief in partner was not found to contribute significantly to any of these regression models.

Correlation with Pro-Vaccine Position

Additionally, anti-vaccine position correlated positively with participants' agreement with the fourth anti-rotation argument (the appeal to personal responsibility) $r(65) = .32$, $p = .010$, suggesting that the intended mimicry of anti-vaccine rhetoric was at least somewhat successful.

Discussion

Contrary to predictions, an encounter with an untrustworthy actor did not, in itself, promote agreement with the anti-tire-rotation arguments. However, and perhaps more importantly, the findings concerning general agreement with the anti-rotation message indicate that while a distrust-provoking experience may make a person a bit more *skeptical*, and a disposition toward distrust of an institution such as the medical establishment does not, by itself, predict favorable responses to criticisms of the status quo, when they are *combined*, these constructs can interact in a very interesting manner – those low in values-based medical mistrust became less favorable to the essay attacking the standard practice of tire rotation, whereas those higher than average became *more* favorable to it after the same distrust-provoking experience. This finding seems to suggest that a person's individual tendency toward suspicious thinking determines whether they respond to a distrust-provoking experience with healthy skepticism or a state of receptivity to the kinds of arguments that are often used to undermine confidence in the expert consensus, and promote abstinence from or (often exploitative) alternatives to the best practices advocated by experts and professional organizations.

The finding that participants higher in MMV responded more favorably to the profit-motive argument, regardless of experimental condition, is also notable. This may suggest that

MMV and institutional distrust generally is rooted in the belief that institutions have interests that conflict with one's own.

These results may also suggest that, at least in terms of their relationship to attitudes, distrust-provoking events serve as *triggers* for underlying personal tendencies to respond with different degrees of skepticism vs. paranoia. Regarding general agreement with the anti-rotation essay, MMV did not have a significant effect when participants had a trust-affirming experience with their game partner. It was only when the encounter violated their trust that MMV became a deciding factor in their agreement with the message.

One clear limitation of this study is that MMV was not measured prior to the administration of the main dependent variables of interest. Further research on this topic should measure MMV prior to, and ostensibly separately from, the distrust manipulation and presentation of pro- and anti-rotation arguments in order to more strongly demonstrate the role of dispositional distrust of institutions.

Chapter 14

General Discussion

The preceding set of experiments sought to evaluate how the cognitive processing of messages may differ in response to a state of diffuse distrust. As discussed in the introductory chapters, prior work has shown that such a state results in non-routine processing. These non-routine processes appear to consist of a heightened sensitivity to contrast, and an increase in cognitive flexibility. Understanding the cognitive effects of distrust from this perspective, I developed a series of experiments seeking to bridge the theoretical gap between the basic cognitive changes associated with distrust and the effect of distrust on attitudes and beliefs.

If distrust increases both sensitivity to contrast and cognitive flexibility, these differences in processing should be observable in the way that people respond to messages and make judgments about people. For those induced to distrust, contradictions between messages should not undermine the believability of messages, and exposure to arguments in favor of a given position should prompt a mental search for contrary information, making people consider a wider range of possible explanations for and alternatives to the information being presented.

Before testing these predictions, it was necessary to conduct eight pilot studies in order to determine which methods of inducing distrust would be effective in the populations available to me. Four distrust manipulations were tested – a word-search priming task, a scrambled-sentence priming task, an impostor-detection task (adapted from Schul, Mayo, & Burnstein, 2008), and an economic deception game (adapted from Posten & Mussweiler, 2013) – on two populations – university students and MTurk workers. These manipulations were judged to be successful if they made people report greater state distrust or perform better on puzzles that require outside-the-box thinking. Both priming methods failed to meet this standard. The impostor-detection task

seemed to work better with MTurk workers than with students, and the economic deception game was fairly successful with both populations. The priming methods were abandoned, and the remaining two manipulations were used in the subsequent experiments.

A few different processes appear to be routes by which the basic cognitive features of distrust could influence the interpretation of messages and formation of attitudes; two sets of experiments examine these processes. The first three experiments were designed to address the hypothesis that diffuse distrust promotes a sensitivity to contrast, the fourth and fifth experiments were designed to test the hypothesis that it promotes cognitive flexibility, and the final experiment was designed to test if diffuse distrust makes people more receptive to arguments in the style of the propaganda circulated by anti-vaccination groups and other denialist movements. In each of these experiments, medical mistrust (Shea et al., 2008) was also measured for exploratory purposes, in order to compare and contrast the role of distrusting beliefs with diffuse state distrust.

Contrast Sensitivity

The results did not reveal strong evidence for the hypothesis that diffuse distrust would make people more sensitive to contrasts. In Experiment 1, participants who were asked to identify impostors did not find statements that contrasted with previously-presented statements more believable than did those who rated people's spontaneity. However, participants who were higher in values-based medical mistrust (Shea et al., 2008), which includes items such as "The Health Care System lies to make money," and "The Health Care System experiments on patients without them knowing," had greater acceptance of factually false statements, as long as they did not contradict a previously-presented statement. Although the experiments in this set sought to better understand how diffuse distrust might promote sensitivity to contrast, this seems to be the

most thought-provoking finding – that a preexisting tendency toward a more paranoid form of distrust might affect how receptive people are to certain kinds of messages, and how able they are to suss out untruth.

In Experiment 2, participants who dealt with an untrustworthy partner in an economic game did not make less stereotypical judgments of different professionals, nor did they find greater contrasts between different characters presented together than when presented alone. Posten and Mussweiler (2013) had found that diffuse distrust made characters elicit less stereotypical trait ratings, and that this effect could be eliminated by giving participants a similarity-finding task before they evaluated the character. Their conclusion was that distrust facilitates the detection of contrasts, in that case, the contrast between a person and their most salient social category. Experiment 2 in the present series was an effort to see if the logic of such a finding could extend to a medical professional's perceived competence, which was pertinent to the original motivation behind this research project – to understand what makes a person more susceptible to anti-vaccine propaganda. The failure of Experiment 2 to find such a result could be due to members of professions not being understood to have as much entitativity (Lickel et al., 2000), or the professions themselves simply not being as strongly associated with their relevant traits as more deeply-ingrained social categories such as ethnicity and gender are. That could help to explain why there was no similar effect of diffuse distrust on the trait ratings of the characters presented.

Experiment 3's participants, having gone through the same economic deception manipulation, did not show increased disagreement with strong arguments for comprehensive senior exams, nor increased agreement with an anti-exam argument, contrary to predictions.

They did, however, produce fewer positive responses to *one* of the pro-exam arguments than did those who dealt with a trustworthy partner.

It is possible, especially in Experiments 1 and 3, that increased contrast sensitivity was not detected in response to the distrust manipulation because these experiments involved social judgment tasks, which involve inherently deeper cognitive processing than did the research they were designed to expand upon. These experiments were intended to extend the findings of other researchers that diffuse distrust led to the non-routine cognitive patterns of identifying contrasting adjectives more quickly than congruent adjectives, and judging exemplars as more distinct from their salient social categories (Posten & Mussweiler, 2013; Schul, Mayo, & Burnstein, 2004) to see if sensitivity to contrasting information would extend to the way that propositional messages are interpreted. Thus, the tasks used in Experiments 1 and 3 were necessarily more cognitively demanding than those used in the experiments they were based on – they required participants to interpret and make judgments about the plausibility of statements (Experiment 1) and generate written responses to paragraph-long arguments (Experiment 3). It is possible that diffuse distrust has effects on the processing of information, but it is only observable when the tasks are quick and simple, and ultimately has no observable effect on the slower, more complex task of evaluating actual arguments.

Cognitive Flexibility

Some evidence did support the cognitive flexibility hypothesis, though, however weakly. In Experiment 4, participants who were asked to identify impostors did not produce more solutions to word stems than participants asked to rate the spontaneity of interviewees. Thus, this manipulation did not seem to make participants more creative in the sense of producing solutions that were both appropriate and novel.

However, in Experiment 5, participants who were paired with an untrustworthy partner in an economic game were more verbose in their efforts to reconcile different statements, provided they believed their partner was a real person. It is worth reiterating, however, that partner type did not interact with statement pairing, as was also predicted. The hypothesis was that diffuse distrust would cause participants to produce more elaborate explanations of how conflicting or directly contradicting statements could co-exist, but not for statements that were already compatible. Participants placed in a state of diffuse distrust produced more text regardless of statement pairing. So, though most of the cognitive-flexibility-based hypotheses did not pan out, diffuse distrust did appear to at least make people more creative, or at least more generative, in that they created more text in solving a problem.

Mayer and Mussweiler (2011) found that creativity, particularly the activation of more categories, increases under distrust when the participant was not being observed by other participants. That the same effect was not observed in the form of completed word stems or reconciliation of different statements is surprising, considering participants were most likely completing these studies privately, without even an experimenter in the room waiting for them to finish (the study being conducted over the Internet).

Schul and colleagues (2008) found that participants who were primed to distrust performed better on non-routine matchstick puzzles, suggesting distrust makes people more cognitively flexible. The present series of experiments did not directly try to replicate this effect, outside of the pilot studies. The failure to replicate Schul et al.'s (2008) effect in my pilot studies could possibly be due to the computer-mediated context (as opposed to real matchsticks on a table), or because of differences in distrust inductions—they used exposure to untrustworthy faces. Perhaps even cultural differences between Americans and Israeli university students

accounts for my inability to find effects son matchstick puzzles. In any case, the fact that my subsequent experiments failed to even conceptually replicate this cognitive flexibility effect is difficult to blame on any one specific procedural difference.

Friesen and Sinclair (2011) found that distrust-primed participants activated both "Black" and "Doctor" categories after watching a video of a Black doctor, whereas participants who were not so primed activated neither. Friesen and Sinclair used a scrambled sentence priming protocol, which I found to be ineffective in my own pilot studies. I did not use activation of social category (or moral valence) as a dependent variable. I did explore category activation through different means (word stem completions) in Experiment 4, but did not find the cognitive flexibility effects I expected. It is possible that Friesen and Sinclair's stimuli, which required participants to make social judgments, were more relevant to the distrust prime in a way that the simplistic stimuli in my experiments were not. Weingarten and colleagues (2016) recently found that priming can be effective when prime and target are directly relevant.

Agreement With Anti-Vaccine-Style Rhetoric

This inquiry into the cognition of distrust was initially motivated by a desire to understand why anti-vaccination propaganda is convincing to people, but the direct cognitive effects of distrust were rather unpredictable in this dissertation. On the other hand, in Experiment 6, where people were presented with the rhetorical style of anti-vaccination arguments, preexisting distrusting (or trusting) beliefs about the world interacted with the presence or absence of diffuse distrust in a fairly intuitive manner. Participants higher in medical mistrust were more favorable toward the novel, anti-tire-rotation arguments after encountering an untrustworthy partner, while those lower on the dispositional measure agreed with them considerably less. Of particular interest is that this effect only seems to happen after the

encounter with an untrustworthy partner. Participants who encountered a trustworthy partner agreed with the anti-rotation messages to about the same degree regardless of their medical mistrust. This seems to imply that a person's beliefs about the world, such as those about the trustworthiness of institutions, influence how they react after they are treated in a deceptive way. Higher MMV score appears to be an indicator of a person's vulnerability to a more paranoid way of thinking that, when activated by a deception attempt, makes them find arguments in the style often found in anti-vaccine propaganda more convincing than they otherwise would be.

People who generally see the world as a place where they can expect to be dealt with honestly (low MMV) seemed to respond to a deception attempt by becoming more skeptical in Experiment 6. They appeared to evaluate arguments more critically and rely more on established experts and institutions.

According to my results, people who are already disposed to feel oppressed or unable to trust others seem to react to deception attempts by becoming more receptive to rhetoric that tells them to be suspicious of others, perhaps particularly institutions that uphold or represent the status quo. Though this response has some of the trappings of a skeptical approach, the tendency seems to be less one of careful consideration of the argument at hand, and more one of contrarian obstinacy. Future research looking to replicate or advance this finding should take care to reproduce my choice to use actionable persuasive messages – cases in which beliefs could be consequential – and not shallow true-or-false statements that lack important behavioral consequences for a participant.

State and Trait Distrust

One important question that arose during the course of this investigation was that of the difference between state distrust and trait distrust. My hypotheses about how manipulations of

participants' state distrust would impact participants' subsequent performance or judgment were generally not supported in this series of experiments, except for in Experiment 6, which found that an interaction with an untrustworthy game partner amplified the relationship between MMV (a dispositional variable) and agreement with an anti-tire-rotation essay. Experiment 1 provided no evidence that manipulated state distrust had an effect on the judged plausibility of statements regardless of how they related to previous statements, but did find that false statements were more believable to participants higher in MMV, as long as those statements were not in contradiction with a previously-presented statement. Given these results, it seems that at least in the context of the present experiments (Americans, online studies), dispositional measures such as MMV are a more reliable means of studying distrust. The results of Experiment 6 also complicate the way we understand distrust manipulations – participants higher in MMV found the anti-tire-rotation message more believable, and those lower in MMV found it less believable – participants' pre-existing beliefs may have important effects on how they respond to situations that are meant to activate distrust. My speculation is that distrust in institutions comes with a general receptiveness to more paranoid types of reasoning, but future research is needed to unpack the underlying mechanisms of this relationship.

Values-Based Medical Mistrust. It is my belief that MMV is useful in a way that the distrust measures used in the pilot studies are not, because it asks about particular concrete fears that a person may have when they are in a position of vulnerability or dependence on a powerful institution. Other measures, such as the IPIP Mistrust scale (Goldberg et al., 2006) or the Trust Questionnaire (Yamagishi & Yamagishi, 1994) orient themselves toward the respondent's beliefs about the general trustworthiness of other people (e.g. "Most people are basically honest," "I feel like people often are out to get something from me") and their own tendency to trust

others (e.g. “I am pretty trusting of others’ motives,” “I am trustful.”), whereas the MMV queries their beliefs about concrete events that could possibly happen to them (e.g. “The Health Care System experiments on patients without them knowing,” “The Health Care System covers up its mistakes.”). This – fearful beliefs stemming from one’s vulnerability to others – is the conception of distrust that is most relevant to the basic philosophical basis of the present project (Ullmann-Margalit, 2004), and is most relevant to how people relate to major institutions that have power over them, such as their employers, medical care providers, banks, and government agencies.

Given the findings of Experiment 6, it appears likely that distrust in the medical system can be used as a proxy for broader institutional distrust – MMV had a predictive effect on agreement with an anti-tire-rotation message, suggesting that a person who is disposed to believe that other people or systems will abuse power over them will also be likely to hold these attitudes across contexts. Ideally, for future research, it would be best to have a broad measure of values-based institutional distrust that addressed explicit beliefs about what harm those institutions may do to them. Such a measure would go beyond the medical establishment and address individuals’ fears about a wide range of powerful social institutions.

Uncertainty and Personal Control

Participants who reported higher uncertainty and lack of personal control produced more words to reconcile contradicting and conflicting pairs of statements and viewing contradicting statements to be more compatible. However, these preexisting tendencies did not predict agreement with the tire rotation essays in Experiment 6. This suggests that uncertainty is not enough to produce the paranoid type of response to anti-vaccine-style rhetoric that we see when participants high in dispositional distrust are provoked by an encounter with an untrustworthy

person. It seems that uncertainty leads people to engage in more extensive processing as demonstrated by Pittman and D'Agostino (1985, 1989; Pittman & Pittman, 1980), but does not inherently make a person more suspicious of another's motives.

Methodological Insights

In pilot testing, the priming methods used in this dissertation – word searches and scrambled sentences – were not reliably effective in changing self-reported trust or distrust measures, performance on routine/nonroutine puzzles, or making similarity judgments. As a result, they were not used in the subsequent experiments. A recent meta-analysis (Weingarten et al., 2016) found that priming is strongest when it activates a goal and when the outcome measure is directly (not indirectly or metaphorically) connected to the content of the prime. The primes in this study were not designed to activate behavioral goals, such as performance on a test or increased frequency of a target behavior. These studies used primes in order to influence participants' reported attitudes and judgments, which were specifically excluded from Weingarten et al.'s analysis. The attitudinal measures of trust and distrust could be considered to be directly related, but the matchstick puzzles and similarity judgments are clearly not, so according to this understanding, the priming effects could be expected to short-lived and somewhat weak. The question remains how other researchers have observed significant effects of primed distrust, but based on the research contained in this document, which is some of very little research of this kind conducted in English, I would not recommend using semantic primes as a means of manipulating distrust, at least in studies conducted online.

The imposter detection paradigm was somewhat effective in pilot tests, but ultimately did not affect the believability of statements in Experiment 1, or the amount of word stem

completions in Experiment 4. This outcome could either be due to the ineffectiveness of the manipulation itself or the incorrectness of the main hypotheses of those experiments.

The economic deception game was fairly effective in influencing responses to various measures in pilot testing, and seemed to affect agreement with the two tire-rotation essays in Experiment 6, but not affect the cognitive process measures of the other experiments in which it was used. Belief in partner did interact with the manipulated variables in Experiment 3, but was not found to be a factor in the case of Experiment 6.

Considering the outcomes in this series of experiments together, it appears that the economic deception game was the most reliable means of experimentally manipulating distrust. Researchers who use this type of manipulation should include questions about partner belief, and should include responses to it in their analysis. Belief had an effect on the target variable in two out of four studies in this series, so it may be worthwhile going forward to investigate what experimental contexts are more or less sensitive to partner belief. They should also accompany this manipulation with a dispositional measure of distrust, such as MMV. Dispositional measures of distrust and experimentally-manipulated distrust appear to be associated with different types of outcomes and may interact with each other. They should not be treated as if they are interchangeable.

General Limitations

There are some limitations to the design of the studies in this document that suggest some directions for future research into the cognitive and attitudinal consequences of distrust.

Dispositional Measures. The dispositional measures in all of these studies were administered *after* the main outcome measure. Because of this they could have been influenced by the main outcome measures rather than moderated them. The studies were designed this way

because these measures were merely exploratory, and thus were included after the critical outcome measures, out of a desire not to contaminate the manipulation by exposing people to the concept of distrust beforehand. Future work in this line of research should pre-test participants for dispositional distrust, particularly values-based medical mistrust in order to bypass this limitation.

Internet-Based Studies. All experiments in this series were conducted online. The most important issue with this choice is that one cannot know if or how participants would have responded differently to the manipulations or measures if they had been administered inside the more tightly-controlled environment of the laboratory. Though participants completed attention checks throughout the experiments, with participants eliminated if they failed them or finished the experiment too quickly or slowly, it is possible that participants were not paying as much attention to the manipulation or the measures as they would have paid in the presence of a human experimenter.

Future research should include some in-laboratory experiments in order to address this unknown. At the same time, it is important to consider that, increasingly, the Internet is becoming the primary way that many people communicate with others, look for information, and even make important real-life decisions and commitments. Despite lacking the control of the laboratory, online studies can be conducted in the very same physical environment in which the participants work, find dating partners, talk with their friends, and shop. In that way, data collected in online studies may represent more relevant responses as to how people are engaging in such behaviors-- responses that correspond more closely to what their real-life responses and decisions would be than would being in a lab.

Implications and Future Directions

That people may respond differently to detected deception attempts according on their beliefs about, or perhaps experiences in the world, could have some important implications for how we understand the way people respond to different styles of persuasive messages. And the fact that such circumstances divide people into responding with a fairly strong rejection of, or a fairly strong agreement with the paranoid rhetorical style of anti-vaccination activists (Experiment 6) seems to track rather well with other political camps that stress the danger of outsiders, deviousness of institutions, and the value of self-reliance and self-interest.

It seems possible that the rising worldwide nationalist fever can in part be explained by growing belief that the people cannot trust their institutions to act in their best interest. At the risk of overextending my findings here, I think they can be applied to our current situation.

With public distrust rising, encouraged by institutional failures both real and concocted by powerful propaganda outlets, the only thing needed for a mass movement based on a paranoid worldview would be for enough people to have some experience of detecting a deception attempt. Just as in the experiments presented here, such an experience can be readily manufactured. From documentaries like “Loose Change,” and “Zeitgeist,” which manipulatively walked their audience through 9/11 conspiracy theories, to “Gamer Gate,” a movement among video gamers filled with misogynistic and anti-media ‘exposé’ videos, there is an entire genre of internet videos to tell the already-suspicious that someone has tricked them in some horrible way.

Returning to the issue of the anti-vaccination movement, one can imagine that for someone who already has distrusting beliefs about the healthcare system, seeing the videos circulating on the internet, or hearing stories from celebrities and politicians about children

experiencing neurological harm in response to vaccination could perform a similar function to being tricked in the economic deception game, and make an already-distrusting person more receptive to the line of arguments that anti-vaccination advocates tend to use.

Of course, in order to confirm and determine how far this interpretation can be extended, further research is necessary. Many questions remain – is this interaction with diffuse distrust unique to values-based distrust in the medical system? Does the distrust need to be directed at an institution? Is it necessary for the deception attempt to be personal and directly experienced, or could a similar effect be achieved by simply explaining to someone how themselves or their group has been deceived?

Furthermore, it seems that it may be fruitful to explore what kinds of life experiences correlate with values-based distrust in the healthcare system. One obvious place to look would be a person's past experiences with the healthcare system, but it may also be worthwhile to see if victims of abuse or other forms of trauma are more likely to endorse these beliefs. Consumption of conspiracy-peddling media could be another predictor of distrust in the medical system or institutions generally.

References

Akehurst, L., Köhnken, G., Vrij, A., & Bull, R. (1996). Lay persons' and police officers' beliefs regarding deceptive behaviour. *Applied Cognitive Psychology*, *10*(6), 461–471. https://doi.org/10.1002/(SICI)1099-0720(199612)10:6<461::AID-ACP413>3.0.CO;2-2

Altemeyer, B. (2006) *The Authoritarians*. Retrieved from http://home.cc.umanitoba.ca/~altemey/ on 8/5/2013.

Bargh, J. A., Chen, M., & Burrows, L. (1996). Automaticity of social behavior: Direct effects of trait construct and stereotype activation on action. *Journal of personality and social psychology*, *71*(2), 230.

Bargh, J. A., Gollwitzer, P. M., Lee-Chai, A., Barndollar, K., & Trötschel, R. (2001). The automated will: nonconscious activation and pursuit of behavioral goals. *Journal of personality and social psychology*, 81(6), 1014.

Benin, A. L., Wisler-Scher, D. J., Shapiro, E. D., & Holmboe, E. S. (2005). Qualitative analysis of mothers' decision-making about vaccines for infants: The importance of trust. *Pediatrics*, *117*(5), 1532–1541.

Bohannon, J. (2014). Replication effort provokes praise—and 'bullying'charges. *Science*, *344*(6186), 788-789.

Cacioppo, J. T., & Petty, R. E. (1984). the Elaboration Likelihood Model of Persuasion. *Advances in Consumer Research*, *11*(1), 673–675. https://doi.org/10.1558/ijsll.v14i2.309

Cialdini, R. (1996). The triple tumor structure of organizational behavior. In D. M. Massick & A. E. Tenbrunsel (Eds.), *Codes of Conduct: Behavioral Research into Business Ethics* (pp. 44–58).

Conchie, S. M., Taylor, P. J., & Charlton, A. (2011). Trust and distrust in safety leadership: Mirror reflections?. *Safety science*, *49*(8), 1208-1214.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology*, *56*(1), 5–18. https://doi.org/10.1037/0022-3514.56.1.5

Dijksterhuis, A., Aarts, H., Bargh, J. A., & van Knippenberg, A. (2000). On the relation between associative strength and automatic behavior. *Journal of Experimental Social Psychology*, *36*(5), 531-544.

Downey, L., Tyree, P. T., Huebner, C. E., & Lafferty, W. E. (2010). Pediatric vaccination and vaccine-preventable disease acquisition: Associations with care by complementary and alternative medicine providers. *Maternal and Child Health Journal*, *14*(6), 922–930. https://doi.org/10.1007/s10995-009-0519-5

Eagly, A. H., Chaiken, S., & Wood, W. (1981). An attribution analysis of persuasion. *New directions in attribution research*, *3*, 37-62.

Edgell, P., Gerteis, J., & Hartmann, D. (2006). Atheists as "other": Moral boundaries and cultural membership in American society. *American Sociological Review*, *71*(2), 211–234. https://doi.org/10.1177/000312240607100203

Fein, S., McCloskey, A. L., & Tomlinson, T. M. (1997). Can the jury disregard that information? The use of suspicion to reduce the prejudicial effects of pretrial publicity and inadmissible testimony. *Personality and Social Psychology Bulletin*, *23*(11), 1215–1226. https://doi.org/10.1177/01461672972311008

Friesen, J., & Sinclair, L., (2011). Distrust and simultaneous activation of multiple categories. *Social Psychological and Personality Science*, *2*(1), 112-118.

Gervais, W. M. (2011). Finding the faithless: perceived atheist prevalence reduces anti-atheist prejudice. *Personality & Social Psychology Bulletin*, *37*(4), 543–56. Retrieved from http://www.ncbi.nlm.nih.gov/pubmed/21343437

Gervais, W. M., & Norenzayan, A. (2012). Reminders of secular authority reduce believers' distrust of atheists. *Psychological Science*, *23*(5), 483–91. https://doi.org/10.1177/0956797611429711

Gilbert, D. T. (1991). How mental systems believe. *American Psychologist, 46*(2), 107-119. doi:10.1037/0003-066X.46.2.107

Goertzel, T., (1994). Belief in conspiracy theories. *Political Psychology, 15*(4), 731-742.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84-96.

Harris, C. R., Coburn, N., Rohrer, D., & Pashler, H. (2013). Two failures to replicate high-performance-goal priming effects. *PloS one*, 8(8), e72467.

Harris, P. L. (2012). *Trusting what you're told: How children learn from others*. Harvard University Press.

Hirsh, E. D. (1988). *Cultural literacy: What every American needs to know.* Boston: Houghton Mifflin.

Jacques, J. T., & Kristensson, P. O. (2013, March). Crowdsourcing a hit: measuring workers' pre-task interactions on microtask markets. In *First AAAI Conference on Human Computation and Crowdsourcing*.

Julka, D. L., & Marsh, K. L. (1993). Attitude functions and the matching hypothesis: Are messages which match an aroused need for information more effective. *annual first-year presentation at the University of Notre Dame*.

Kaufmann, N., Schulze, T., & Veit, D. (2011, August). More than fun and money. Worker Motivation in Crowdsourcing-A Study on Mechanical Turk. In *AMCIS* (Vol. 11, pp. 1-11).

Kawakami, K., Young, H., & Dovidio, J. F. (2002). Automatic Stereotyping: Category, Trait, and Behavioral Activations. *Personality and Social Psychology Bulletin, 28*(1), 3–15. https://doi.org/10.1177/0146167202281001

Kay, A. C., Shepherd, S., Blatz, C. W., Chua, S. N., & Galinsky, A. D. (2010). For God (or) country: the hydraulic relation between government instability and belief in religious sources of control. *Journal of Personality and Social Psychology*, *99*(5), 725–739. https://doi.org/10.1037/a0021140

Kittur, A., Chi, E. H., & Suh, B. (2008, April). Crowdsourcing user studies with Mechanical Turk. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 453-456). ACM.

Kramer, R. M. (1994). The sinister attribution error: Paranoid cognition and collective distrust in organizations. *Motivation and Emotion, 18*(2), 199–230. https://doi.org/10.1007/BF02249399

Kramer, R. M. (1999). TRUST AND DISTRUST IN ORGANIZATIONS : Emerging Perspectives , Enduring Questions.

Lewandowsky, S., Ecker, U. K., Seifert, C. M., Schwarz, N., & Cook, J. (2012). Misinformation and its correction: continued influence and successful debiasing. *Psychological Science in the Public Interest*, *13*(3), 106-131.

Lewandowsky, S., Oberauer, K., & Gignac, G. E. (2013). NASA faked the moon landing—therefore, (climate) science is a hoax: An anatomy of the motivated rejection of science. *Psychological Science*, *24*(5), 622-633. doi:10.1177/0956797612457686

Marsh, K. L., & Weary, G. (1989). Depression and attributional complexity. *Personality and Social Psychology Bulletin*, *15*(3), 325-336.

Marsh, K. L., & Weary, G. (1994). Severity of depression and responsiveness to attributional information. *Journal of Social and Clinical Psychology*, *13*(1), 15–32.

Mayer, J., & Mussweiler, T. (2011). Suspicious spirits, flexible minds: When distrust enhances creativity. *Journal of Personality and Social Psychology 101*(6), 1262-1277.

Mayo, R., Alfasi, D., & Schwarz, N. (2014). Distrust and the positive test heuristic: dispositional and situated social distrust improves performance on the Wason rule discovery task. *Journal of Experimental Psychology. General*, *143*(3), 985–90. https://doi.org/10.1037/a0035127

Norenzayan, A., & Gervais, W. M. (2015). Secular Rule of Law Erodes Believers' Political Intolerance of Atheists. *Religion, Brain & Behavior*, *5*(1), 3–14. https://doi.org/10.1017/CBO9781107415324.004

Omer, S. B., Richards, J. L., Ward, Michelle, A. B., & Bednarczyk, R. A. (2012). Vaccination policies and rates of exemption from immunization 2005-2011. *New England Journal of Medicine*, *367*(12), 1170–1171. https://doi.org/10.1056/NEJMc1208306

Ou, C. X., & Sia, C. L. (2010). Consumer trust and distrust: An issue of website design. *International Journal of Human-Computer Studies, 68*(12), 913-934.

Pashler, H., Coburn, N., & Harris, C. R. (2012). Priming of social distance? Failure to replicate effects on social and food judgments. *PloS one,* 7(8), e42510.

Pashler, H., Rohrer, D., & Harris, C. R. (2013). Can the goal of honesty be primed?. *Journal of Experimental Social Psychology*, *49*(6), 959-964.

Petty, R. E., & Cacioppo, J. T. (1979). Issue involvement can increase or decrease persuasion by enhancing message-relevant cognitive responses. *Journal of personality and social psychology*, *37*(10), 1915.

Petty, R.E., & Cacioppo, J.T. (1986). The elaboration likelihood model of persuasion. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology*(Vol. 19, pp. 123-205). New York, Academic Press.

Pittman, T. S., & D'Agostino, P. R. (1985). Motivation and attribution: The effects of control deprivation on subsequent information processing. In J. H. Harvey & G. Weary (Eds.), *Attribution: basic issues and applications* (pp. 117–141). Academic Press. https://doi.org/10.4319/lo.2013.58.2.0489

Pittman, T. S., & D'Agostino, P. R. (1989). Motivation and cognition: Control deprivation and the nature of subsequent information processing. *Journal of Experimental Social Psychology*, *25*(1), 465–480. https://doi.org/10.4319/lo.2013.58.2.0489

Pittman, T. S., & Pittman, N. L. (1980). Deprivation of control and the attribution process. *Journal of Personality and Social Psychology*, *39*(3), 377–389. https://doi.org/10.1037/0022-3514.39.3.377

Posten, A. C., Ockenfels, A., & Mussweiler, T. (2014). How activating cognitive content shapes trust: A subliminal priming study. *Journal of Economic Psychology*, *41*, 12–19.

Posten, A., & Mussweiler, T. (2013). When Distrust Frees Your Mind: The Stereotype-Reducing Effects of Distrust. *Journal Of Personality And Social Psychology*, doi:10.1037/a0033170

Pratto, F., Çidam, A., Stewart, A. L., Zeineddine, F., Aranda, M., Aiello, A., & ... Henkel, K. E. (2013). Social dominance in context and in individuals: Contextual moderation of robust effects of social dominance orientation in 15 languages and 20 countries. *Social Psychological And Personality Science*, *4*(5), 587-599. doi:10.1177/1948550612473663

Raudenbush, S., Bryk, A., & Congdon, R. (2013). HLM 7.01 for Windows [Hierarchical linear and nonlinear modeling software]. *Los Angeles: Multivariate Software*.

Salmon, D. A., Haber, M., Gangarosa, E. J., Phillips, L., Smith, N. J., & Chen, R. T. (1999). Health consequences of religious and philosophical exemptions from immunization laws. *Jama*, *281*(1), 47–54.

Schul, Y., & Peri, N. (2015). Influences of Distrust (and Trust) on Decision Making. *Social Cognition*, *33*(5), 414–435. https://doi.org/http://dx.doi.org/101521soco2015335414

Schul, Y., Burnstein, E., & Bardi, A. (1996). Dealing with Deceptions that are Difficult to Detect: Encoding and Judgment as a Function of Preparing to Receive Invalid Information. *Journal of Experimental Social Psychology*, *32*(3), 228–253. https://doi.org/http://dx.doi.org/10.1006/jesp.1996.0011

Schul, Y., Mayo, R., & Burnstein, E., (2008) The value of distrust. *Journal of Experimental Social Psychology, 44*, 1293-1302.

Schul, Y., Mayo, R., Burnstein, E., & Yahalom, N. (2007). How people cope with uncertainty due to chance or deception. *Journal of Experimental Social Psychology*, *43*(1), 91–103. https://doi.org/10.1016/j.jesp.2006.02.015

Schul, Y., Mayo, R., Burnstein, E., & Yahalom, N. (2007). How people cope with uncertainty due to chance or deception. *Journal of Experimental Social Psychology*, *43*(1), 91–103. https://doi.org/10.1016/j.jesp.2006.02.015

Schul, Y., Mayo, R., Burnstein, E., (2004). Encoding under trust and distrust: The spontaneous activation of incongruent cognitions. *Journal of Personality and Social Psychology*, *86* (5), 668-679.

Shanks, D. R., Newell, B. R., Lee, E. H., Balakrishnan, D., Ekelund, L., Cenac, Z., ... & Moore, C. (2013). Priming intelligent behavior: An elusive phenomenon. *PloS one*, *8*(4), e56515.

Shea, J.A., Micco, E., Dean, L.T., McMurphy, S., Schwartz, J.S., & Armstrong, K.(2008). Development of a revised health care system distrust scale. *Journal of general internal medicine*, *23*(6), 727-732.

Shoff, C., & Yang, T. C. (2012). Untangling the associations among distrust, race, and neighborhood social environment: A social disorganization perspective. *Social science & medicine*, *74*(9), 1342-1352.

Sternberg, R. J., & Lubart, T. I. (1999). The concept of creativity: Prospects and paradigms. *Handbook of creativity*, *1*, 3-15.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207–232. https://doi.org/10.1016/0010-0285(73)90033-9

Twenge, J. M., Campbell, W. K., & Carter, N. T. (2014). Declines in Trust in Others and Confidence in Institutions Among American Adults and Late Adolescents, 1972-2012. *Psychological Science*, 25(10), 1914–1923. https://doi.org/10.1177/0956797614545133

Weary, G., Marsh, K. L., Gleicher, F., & Edwards, J. A. (1993). Depression, Control Motivation and Social Cognition. In G. Weary, F. Gleicher, & K. L. Marsh (Eds.), *Control Motivation and Social Cognition* (pp. 255–287). New York, NY: Springer New York. https://doi.org/10.1007/978-1-4613-8309-3_10

Weingarten, E., Chen, Q., McAdams, M., Yi, J., Hepler, J., & Albarracín, D. (2016). From primed concepts to action: A meta-analysis of the behavioral effects of incidentally presented words. *Psychological Bulletin*, *142*(5), 472–497. http://doi.org/10.1037/bul0000030

Winter, L., & Uleman, J. S. (1984). When are social judgments made? Evidence for the spontaneousness of trait inferences. *Journal of Personality and Social Psychology*, *47*(2), 237–252. https://doi.org/10.1037/h0090437

Wolfe, R.M., Sharp, L.K., & Lipsky, M.S. (2002) Content and design attributes of antivaccination websites. *JAMA*, *287*(24). 3245-3248.

Wood, J., & Berg, P. (2011). Trust in Banks. Gallup, Inc.

Wright, A. G., & Simms, L. J. (2014). On the structure of personality disorder traits: Conjoint analyses of the CAT-PD, PID-5, and NEO-PI-3 trait models. Personality Disorders: Theory, Research, and Treatment, 5(1), 43.

Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion*, 18(2), 129-166.

Yamagishi, T., Satoshi, A., Cho, K., Yumi, I., Li, Y., & Matsumoto, Y. (2015). Two-Component Model of General Trust: Predicting Behavioral Trust from Attitudinal Trus. *Social Cognition*, *33*(5), 436–458. https://doi.org/10.1521/soco.2015.33.5.436

Zebrowitz, L. A. (1997). *Reading faces: Window to the soul?*. Westview Press.

Zimmerman, R. K., Wolfe, R. M., Fox, D. E., Fox, J. R., Nowalk, M., Troy, J. A., & Sharp, L. K. (2005). Vaccine Criticism on the World Wide Web. *Journal Of Medical Internet Research*, *7*(2), doi:10.2196/jmir.7.2.e17

Appendix A

Confidence and Personal Control Measures

From: (Weary, Marsh, Gleicher, & Edwards, 1993)

Scale: 1 (Strongly Disagree) to 5 (Strongly Agree)

Confidence:

At this moment, I feel confident in the correctness of my opinions, judgments, and decisions.

Personal Control:

At this moment, I feel I have sufficient personal control and power over the things that happen in my life.

Appendix B

IPIP Mistrust (Goldberg et al., 2006)

Describe yourself as you feel at this very moment, not necessarily as you feel generally.

Scale: 1 (Very Inaccurate) to 5 (Very Accurate)

1. I feel that others are out to get me
2. I am pretty trusting of others' motives (Reverse-scored)
3. I feel like people often are out to get something from me
4. I suspect hidden motives in others
5. I believe that, sooner or later, people always let you down
6. I believe that people are basically honest and good (Reverse-scored)

Appendix C

Trust for Specific Other Persons and Institutions

(Adapted from Goertzel, 1994)

Scale: 1 (Strongly Disagree) to 5 (Strongly Agree)

1. I feel I can trust my neighbors.
2. I feel I can trust my relatives.
3. I feel I can trust myself.
4. Please answer "Disagree" for this item. (attention check)
5. I feel I can trust science.
6. I feel I can trust my religion.
7. I feel I can trust the police. (Included when stated in methods)
8. I feel I can trust the government (Included when stated in methods)

Appendix D

Trust Questionnaire (Yamagishi & Yamagishi, 1994):

Scale:1 (Strongly Disagree) to 6 (Strongly Agree)

Trust Factor

1. Most people are basically honest.
2. Most people are trustworthy.
3. Most people are basically good and kind.
4. Most people are trustful of others.
5. I am trustful.
6. Most people will respond in kind when they are trusted by others.

Caution Factor

1. No matter what they say, most people inwardly dislike putting themselves out to help others.
2. People are always interested in their own welfare.
3. There are many hypocrites in this society.
4. In this society, one does not need to be constantly afraid of being cheated.
5. One can avoid falling into trouble by assuming that all people have a vicious streak.
6. People usually do not trust others as much as they say they do.

In this society, one has to be alert or someone is likely to take advantage of you.

Appendix E

Matchstick Puzzles

You are about to encounter some mathematical puzzles. You will be presented with eight (8) incorrect equations written with matchsticks.

Your task will be to move one matchstick from one symbol to another in order to make the equation correct.

So, if you are given the problem:

XI = X + VI

You could move one of the matchsticks from the "V" on the right side of the equation over to the left side of the equation, so it would read:

XII = X + II

You will have a text box beneath each equation. Please enter the corrected equation using Roman

numerals.

In case you need a refresher on your Roman numerals, you will be provided with this guide on the next page:

Roman	Arabic
I	1
II	2
II	3
IV	4
V	5
VI	6
VII	7
VIII	8
IX	9
X	10
XI	11
XII	12
XIII	13
XIV	14
XV	15

You will have a time limit of 3 minutes to complete all 8 puzzles on the next page.

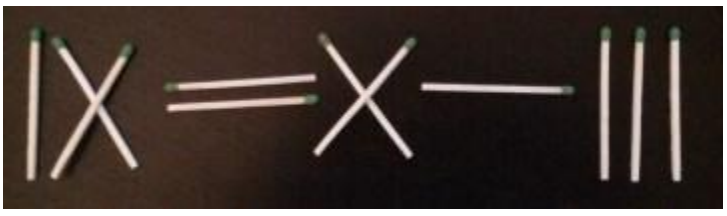
Move one matchstick from one symbol to another in order to make the equation correct. Enter your answer in Roman numerals in the text box below the matchstick equation.

You have 3 minutes to complete as many items as possible.

If you need help reading the Roman numerals, refer to the following figure:

Roman	Arabic
I	1
II	2
II	3
IV	4
V	5
VI	6
VII	7
VIII	8
IX	9
X	10
XI	11
XII	12
XIII	13
XIV	14
XV	15

IX = X - III



Move one matchstick to make the above equation correct. Please type-in your answer using Roman numerals in the text box below.

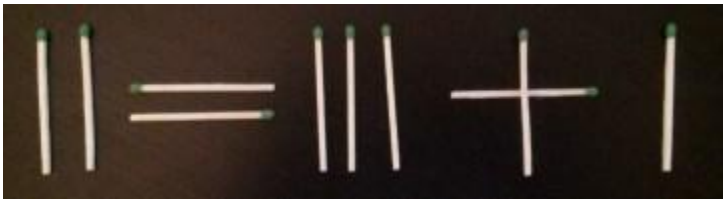
_____

VI = VIII - IV

Move one matchstick to make the above equation correct. Please type-in your answer using Roman numerals in the text box below.

_____

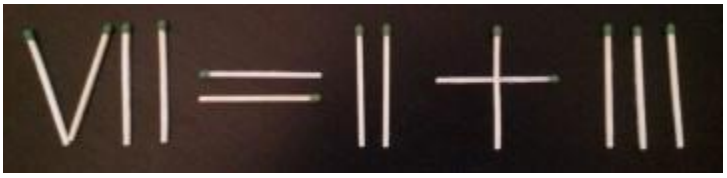
II = III + I



Move one matchstick to make the above equation correct. Please type-in your answer using Roman numerals in the text box below.

_____

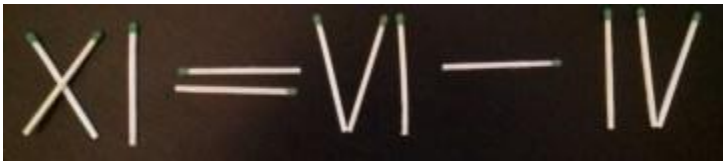
VII = II + III



Move one matchstick to make the above equation correct. Please type-in your answer using Roman numerals in the text box below.

_____

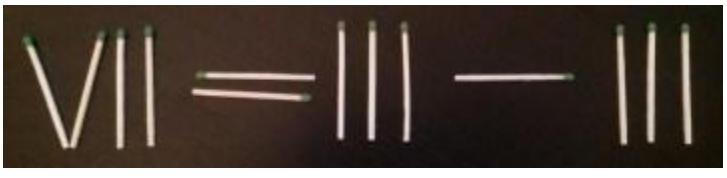
XI = VI - IV



Move one matchstick to make the above equation correct. Please type-in your answer using Roman numerals in the text box below.

_____

VII = III - III



Move one matchstick to make the above equation correct. Please type-in your answer using Roman numerals in the text box below.

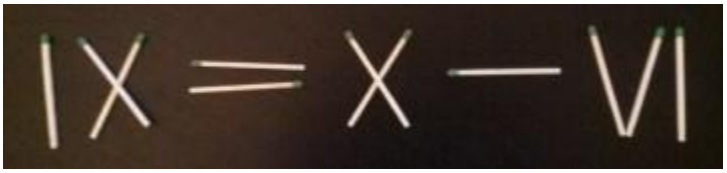
_____

II = VII + IV

Move one matchstick to make the above equation correct. Please type-in your answer using Roman numerals in the text box below.

_____

IX = X - VI



Move one matchstick to make the above equation correct. Please type-in your answer using Roman numerals in the text box below.

_____

That's it for the matchsticks! If you didn't finish all of them, that's fine. It is intended to be difficult.

Please click through to continue.

Appendix F

Similarity Judgments

Scale: 1 (Very Dissimilar) to 6 (Very Similar)

Please judge how similar or dissimilar the following pairs of objects are.

1. Tree and Flower are...
2. Lake and Ocean are...
3. Soda and Water are...
4. Lion and House Cat are...
5. Banana and Cherry are...
6. Laptop and PC are...
7. Glasses and Contacts are...
8. Hat and Turban are...
9. House and Apartment are...
10. Pants and Shorts are...
11. Water and Water are...
12. Mountain and Hill are...
13. Email and Letter are...
14. Eagle and Robin are...
15. Earring and Necklace are...
16. Bread and Chocolate are...

Note: Items were presented one at a time.

Appendix G

IPIP Variety-Seeking (Goldberg et al., 2006)

Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to the other people you know of the same sex as you are, roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in an absolute confidence.

Scale: 1 (Very Inaccurate) to 5 (Very Accurate)

1. Prefer to stick with things I know (Reverse-scored)
2. Dislike changes
3. Love to think up new ways of doing things
4. Like to begin new things
5. Don't like the idea of change (Reverse-scored)
6. Seek adventure
7. Please answer "Neither Accurate Nor Inaccurate" for this item. (Attention check)
8. Enjoy hearing new ideas
9. Like to visit new places
10. Prefer variety to routine
11. Am open to change

Appendix H

IPIP Intellectual Openness (Goldberg et al., 2006)

Describe yourself as you generally are now, not as you wish to be in the future. Describe yourself as you honestly see yourself, in relation to the other people you know of the same sex as you are, roughly your same age. So that you can describe yourself in an honest manner, your responses will be kept in an absolute confidence.

Scale: 1 (Very Inaccurate) to 5 (Very Accurate)

1. Am not interested in abstract ideas (Reverse-scored)
2. Prefer to stick with things that I know (Reverse-scored)
3. Prefer variety to routine
4. Am interested in many things
5. Rarely look for a deeper meaning in things (Reverse-scored)
6. Am open to change
7. Am not interested in theoretical discussions (Reverse-scored)
8. Carry the conversation to a higher level
9. Want to increase my knowledge
10. Try to avoid complex people (Reverse-scored)

Appendix I

Demographic Questions

What is your Age?

What is your gender and/or sex? Please select all that apply.

- ☐ Male (1)
- ☐ Female (2)
- ☐ Female to Male Transsexual (3)
- ☐ Male to Female Transsexual (4)
- ☐ Intersex (5)
- ☐ Other Transgender (6)

How would you describe yourself in terms of race/ethnicity?

Check all that apply.

- ☐ White/Caucasian (1)
- ☐ Black/African American (2)
- ☐ Asian (3)
- ☐ Native Hawaiian/other Pacific Islander (4)
- ☐ Hispanic/Latino (5)
- ☐ Other (please specify) (6) _____

Economically, do you consider yourself:

- ○ Impoverished (1)
- ○ Lower Class (2)
- ○ Lower-Middle Class (3)
- ○ Middle Class (4)
- ○ Upper-Middle Class (5)
- ○ Upper Class (6)
- ○ Super Rich (7)

Appendix J

Position on Childhood Vaccines

Which of the statements below is closest to your own position on childhood vaccines?

- ○ Children should not receive any vaccines at all. (1)
- ○ Children should receive some, but not all recommended vaccines. (2)
- ○ Children should receive all recommended vaccines. (3)

Appendix K

Scrambled Sentences

Instructions: Following are 15 sets of 5 words. For each set of 5 words, remove one word, and re-arrange the remaining 4 words in a way that produces a grammatically-correct sentence.

For example, if the scrambled sentence reads,

"cracked Stella seashells some found"

You might solve it as,

"Stella found some seashells"

Type your answers into the text field below each scrambled sentence.

DISTRUST SET

1	hat	a	Steve	wore	computer	
2	distrust	had	two	parts	Tom	
3	were	there	there	element	some	
4	suspiciously	quickly	ran	they	out	
5	habitual		piece	I	found	a
6	stay	doubt	to	they	had	
7	part	the	had	some	radishes	
8	gasoline		threw	she	the	hook
9	that	more	was	Saturday	fun	
10	questionable	it	was	not	smooth	
11	she	was	very	over	fake	
12	down	did	mistaken	fall	she	
13	was	Frank	never	uncertain	there	
14	tall	embarrassing	it	so	is	
15	John	was	line	in	suspicion	

TRUST SET

1	hat	a	Steve	wore	computer	
2	trusting		had	two	parts	Tom
3	were	there	there	element		some
4	honestly		quicklyran		they	out
5	habitual		piece	I	found	a
6	stay	authentic		to	they	had
7	part	the	had	some	radishes	
8	gasoline		threw	she	the	hook
9	that	more	was	Saturday		fun
10	reliableit		was	not	smooth	
11	she	was	very	over	credible	
12	down	did	sympathetic		fall	she
13	was	Frank	never	certain	there	
14	tall	embarrassing		it	so	is
15	John	was	line	in	dependence	

CONTROL SET

1	hat	a	he	wore	computer	
2	apartment		had	two	parts	he
3	were	there	there	element		some
4	training		quicklyran		they	out
5	habitual		piece	I	found	a
6	stay	elephant		to	they	had
7	part	the	had	some	radishes	
8	gasoline		threw	she	the	hook
9	that	more	was	Saturday		fun
10	oldest	it	was	not	smooth	
11	she	was	there	over	excited	
12	down	did	bookish		fall	she
13	was	one	never	swift	there	
14	tall	embarrassing		it	so	is
15	he	was	line	in	fancy	

Appendix L

Imposter Detection Task Stimuli

You are about to read some statements collected from some athletes at the University of Connecticut about life as a student athlete.

Your task will be to carefully read the responses, and to judge the spontaneity of each respondent.

First, just carefully read each of the responses. After you have read each one, they will be shown to you again for you to make your final decisions.

Respondent 1

What are the most important traits for a successful athlete to have? "Ambition Dedication"

What sport do you play for this school? "Women's Soccer"

What do you enjoy most about your sport? "The team effort – no one person can win the game."

What do you like to do in your free time? "Watch TV and catch up with friends."

If you could change one rule about your sport, what would it be? "Off sides"

What is the hardest part about being a student athlete? "Balancing school and sports and time for myself."

What is your major area of study? "Individualized: Neuroscience"

What is most satisfying about your sport? "Knowing your contribution can help win the game."

Respondent 2

What are the most important traits for a successful athlete to have?
"Determination, patience, passion"

What sport do you play for this school?

"Rowing"

What do you enjoy most about your sport?
"The sunrises and sound of the oarlocks, it's the little things."

What do you like to do in your free time?
"Working on campus at my student job"

If you could change one rule about your sport, what would it be?
"Race time starting rules"

What is the hardest part about being a student athlete?
"Finding the priority between my sport, my major, and sleep"

What is your major area of study?
"Allied Health – DGS"

What is most satisfying about your sport?
"You can never be perfect, you always strive for more"

Respondent 3

What are the most important traits for a successful athlete to have?
"The ability to manage time effectively to accomplish the necessary work in the classroom and on the field/in the court. A passion for playing the game and working hard to improve yourself and your teammates."

What sport do you play for this school? "Basketball"

What do you enjoy most about your sport? "Nobody is perfect in basketball. It is a physical and mental challenge to perform your best on a daily basis and move on to the next play both when you perform well and when you make mistakes."

What do you like to do in your free time? "Watch TV and hang out with friends who I don't get to see much during the season."

If you could change one rule about your sport, what would it be? "Have every jump ball actually be a jump ball instead of a change in possession."

What is the hardest part about being a student athlete? "Balancing time spent between the sport, schoolwork, and relaxing."

What is your major area of study? "Molecular and Cellular Biology"

What is most satisfying about your sport? "Basketball is a sport where there is really no limit to how much you can improve, and sometimes your improvement can be visible over a very short period of time. There is no better thing than realizing that your hard work has paid off and that you have improved as a basketball player because of it."

Respondent 4

What are the most important traits for a successful athlete to have?
"The most important traits for a successful athlete to have is teamwork, perserverence, personable, and hard-working."

What sport do you play for this school?
"Soccer"

What do you enjoy most about your sport?
"The joy the sport gives me. It takes my mind off everything."

What do you like to do in your free time?
"Hang out with my friends."

If you could change one rule about your sport, what would it be?
"The teams we play and travel schedule."

What is the hardest part about being a student athlete?
"Being a good student and a good athlete."

What is your major area of study?
"Business - Economics."

What is most satisfying about your sport?
"Winning w/ my team."

Appendix M

The Economic Deception Game

The Economic Deception Game is presented below, lightly edited from the Qualtrics export. The branching logic is contained in the boxed notes.

Please read these instructions carefully.

When you go to the next page, you will be connected to another participant. Together, the two of you will be deciding how to distribute a maximum of 6 virtual raffle tickets for a $20 Amazon gift card between yourselves.

There are two ways the tickets can be distributed:

- Each Player gets 3 tickets

- 4 tickets for Player 1 and 1 ticket for Player 2

You will be randomly assigned to be either Player 1 or Player 2. Each player will have a unique role in the decision-making process:

- Player 1 will be able to see which option (A or B) represents which distribution, and will recommend one of the options to Player 2.

- Player 2 will not be able to see which option represents which distribution, but will receive Player 1's recommendation and will make the final decision about which option will be chosen.

Q197 Please indicate your understanding of the instructions. Which of these is correct?

- ○ Player 2 knows which option is which. (1)
- ○ Player 1 makes the final decision. (2)
- ○ Player 2 gets a recommendation from Player 1. (3)

Page Break

Display This Question:

If Please indicate your understanding of the instructions. Which of these is correct? != Player 2 gets a recommendation from Player 1.

Q295 Your previous answer was INCORRECT!

Please read these instructions carefully.

When you go to the next page, you will be connected to another participant. Together, the two of you will be deciding how to distribute a maximum of 6 virtual raffle tickets for a $20 Amazon gift card between yourselves.

There are two ways the tickets can be distributed:

- Each Player gets 3 tickets

- 4 tickets for Player 1 and 1 ticket for Player 2

You will be randomly assigned to be either Player 1 or Player 2. Each player will have a unique role in the decision-making process:

- Player 1 will be able to see which option (A or B) represents which distribution, and will recommend one of the options to Player 2.

- Player 2 will not be able to see which option represents which distribution, but will receive Player 1's recommendation and will make the final decision about which option will be chosen.

Please indicate which of the following statements are true:

- ○ Player 1 knows which option is which. (1)
- ○ Player 1 makes the final decision. (2)
- ○ Player 2 gets a recommendation from Player 1. (3)

End of Block: Participant pool info

Start of Block: Computer Simulation

Q164 Please wait while the computer connects to another participant . . .

End of Block: Computer Simulation

Start of Block: Computer simulation 2

Q175 Please wait while the computer assigns roles. . .

End of Block: Computer simulation 2

Start of Block: Trust condition

Q176 You are assigned to Player 2.

Player 1 is evaluating the options....

Page Break

Q201 You are assigned to Player 2.

Player 1 suggests you pick option A.

"that gives each of us 3"

Q177 Please choose an option.

- A (1)
- B (2)

End of Block: Trust condition

Start of Block: Distrust Condition

Q182 You are assigned to Player 2.

Player 1 is evaluating the options....

Page Break

Q203 You are assigned to Player 2.

Player 1 suggests you pick option A.

"that gives each of us 3"

Q183 Please choose an option.

- A (1)
- B (2)

End of Block: Distrust Condition

Start of Block: Trust outcome A

Display This Question:

If Please choose an option. = A

Q179
The Selected Distribution is:

You: 3 TICKETS
Player 1: 3 TICKETS

If you had gone against your partner's advice, you would have received 1 ticket, and they would have received 4.

End of Block: Trust outcome A

Start of Block: Trust outcome B

Display This Question:

If Please choose an option. = B

Q181
The Selected Distribution is:

You: 1 TICKET
Player 1: 4 TICKETS

If you had taken Player 1's advice, you would each have been awarded 3 tickets.

End of Block: Trust outcome B

Start of Block: Distrust outcome A

Display This Question:

If Please choose an option. = A

Q184
The Selected Distribution is:

You: 1 TICKET
Player 1: 4 TICKETS

If you had gone against your partner's advice, you would each have been awarded 3 tickets.

End of Block: Distrust outcome A

Start of Block: Distrust outcome B

Display This Question:

If Please choose an option. = B

Q185
The Selected Distribution is:

You: 3 TICKETS
Player 1: 3 TICKETS

If you had taken Player 1's advice, you would have received 1 ticket, and they would have received 4.

End of Block: Distrust outcome B

Start of Block: Block 14

Q195 We have made note of how many tickets you have been awarded. The drawing will happen once data collection is complete. If you are the winner, you will be notified via your UCONN email.

Please continue on and answer the following questions honestly and as accurately as possible.

End of Block: Block 14

Questions about Player 1, included at the end of the session:

Think back to the very beginning of the study.

What do you think about your interaction partner (Player 1)?

_____

Again, think back to the beginning of the study. Due to the limited interaction between partners in this study, some participants may not believe that they are matched up with a real person.

Do you think your partner believed you were a real participant?

- ○ Yes (1)
- ○ No (2)

Do you believe your partner was a real participant?

- ○ Yes (1)
- ○ No (2)

Page Break

In reality, your partner was not a real person, but was simulated by the computer. How surprised are you at this information?

- ○ Extremely Unsurprised (1)
- ○ Somewhat Unsurprised (2)
- ○ Slightly Unsurprised (3)
- ○ Slightly Surprised (4)
- ○ Somewhat Surprised (5)
- ○ Extremely surprised (6)

Appendix N

Experiment 1 Statements

Statement sets were presented as if they were one set. They are presented here in order of presentation, labeled by set, truth/falsity, and pairing type.

Set 1

True/False	Pairing Type	Statement
False	Contradicting	Jack London's first major novel was *Billy Budd*.
True	Irrelevant	England was the European nation primarily responsible for exploring and settling the east coast of the United States
True	Agreeing	Mach 1 refers to the speed of sound
True	Contradicting	Mount Everest is located in Nepal
False	Irrelevant	Shakespeare said, "O, what a tangled web we weave, when first we practice to deceive!"
True	Agreeing	Women have been able to vote in Federal elections for less than 100 years
False	Irrelevant	The Roman numeral for 50 is C
True	Agreeing	The American Civil War took just over four years to come to an end.
True	Contradicting	The purpose of The Federalist papers was to gain ratification of the United States Constitution
True	Contradicting	The ruins of Troy are located in Turkey
False	Agreeing	*Julius Caesar* by Shakespeare is a play about Caesar's love affair with Cleopatra
True	Irrelevant	10 degrees Celcius is equivalent to 50 degrees Fahrenheit

False	Contradicting	In 1939, the United States was pulled into WWII by a surprise attack at Pearl Harbor.
False	Irrelevant	The Old Testament said, "Neither a borrower or a lender be; For loan oft loses both itself and friend, And borrowing dulls the edge of husbandry"
False	Agreeing	Sydney is the capital of Australia

Set 2

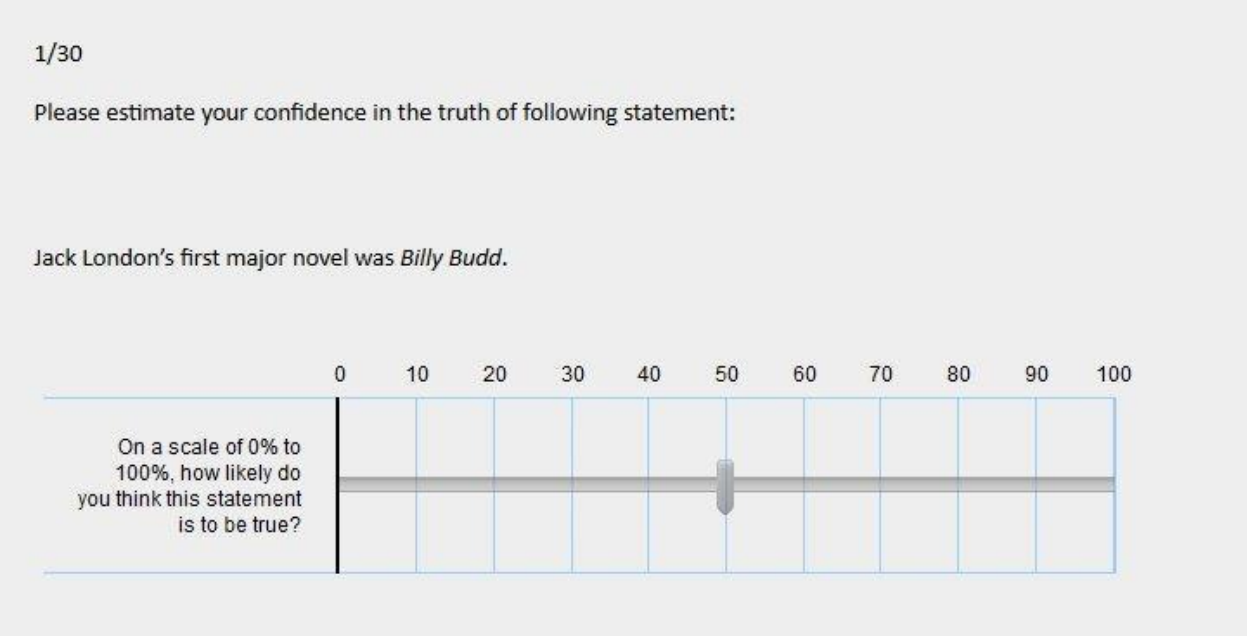
True/False	Pairing Type	Statement
False	Contradicting	Every year, Greece welcomes thousands of tourists eager to see the ruins of Troy.
False	Irrelevant	Lincoln's presidential term lasted from 1860 to 1865
True	Agreeing	Cleopatra and Julius Caesar had a legendary affair centuries ago.
True	Contradicting	The attack on Pearl Harbor occurred in 1941
False	Irrelevant	The Consitution was written between 1771 and 1780
True	Agreeing	Women were given the right to vote in 1920
False	Irrelevant	The Jewish New Year is referred to as Yom Kippur
True	Agreeing	A vehicle must accelerate past Mach 1 to break the sound barrier
False	Contradicting	Mount Everest is one of Bangladesh’s most famous landmarks
False	Contradicting	The American Revolution was able to obtain foreign approval due to the success of the Federalist Papers.
False	Agreeing	The Civil War lasted from 1856 to 1860

True	Contradicting	*Billy Budd, Benito Cereno, and Batleby the Scrivener* were written by Herman Melville
True	Irrelevant	In the 1984 democratic primaries, Walter Mondale ran against Gary Hart
False	Agreeing	Australia's parliament building can be found in Sydney.

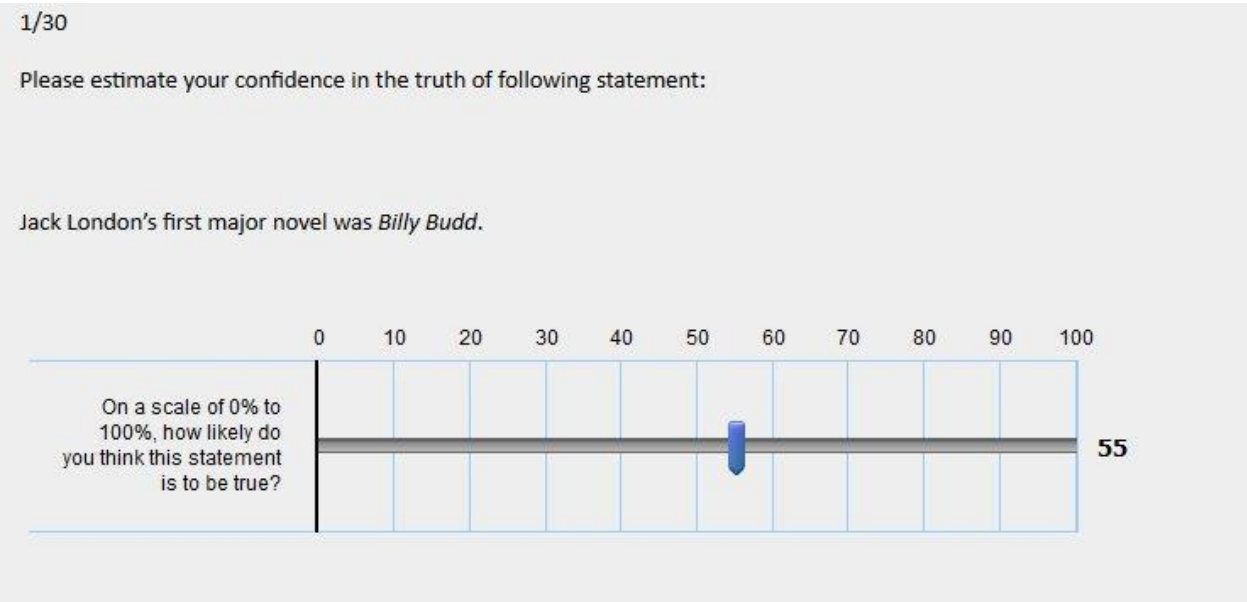
Appendix O

Plausibility Rating Scales

The default position of the slider, resulting in "missing data:"



The slider, when a value is selected correctly:



Appendix P

Revised Health Care System Distrust Scale (Shea, Micco, Dean, et al., 2008)

Items for the competence factor are indicated with MMC. Items for the values factor are indicated with MMV.

Scale: 1 (Strongly disagree) to 5 (Strongly agree)

1. The Health Care System does its best to make patients' health better (MMC, Reverse-Scored)
2. The Health Care System covers up its mistakes (MMV)
3. Patients receive high quality medical care from the Health Care System (MMC, Reverse-Scored)
4. The Health Care System makes too many mistakes (MMC)
5. The Health Care System puts making money above patients' needs (MMV)
6. The Health Care System gives excellent medical care (MMC, Reverse-Scored)
7. Patients get the same medical treatment from the Health Care System, no matter what the patient's race or ethnicity (MMV, Reverse-Scored)
8. The Health Care System lies to make money (MMV)
9. The Health Care System experiments on patients without them knowing (MMV)

Appendix Q

Short Social Dominance Orientation (SSDO) Questionnaire (Pratto, Çidam, Stewart, et al., 2013)

Scale: 1 (Strongly Disfavor) to (Strongly Favor)

1. In setting priorities, we must consider all groups. (Reverse-Scored)
2. We should not push for group equality.
3. Group equality should be our ideal. (Reverse-Scored)
4. Superior groups should dominate inferior groups.

Appendix R

Right Wing Authoritarianism (RWA) Quesionnaire (Altemeyer, 2006)

Scale: 1 (Very Strongly Disagree) to 9 (Very Strongly Agree)

1. The established authorities generally turn out to be right about things, while the radicals and protestors are usually just “loud mouths” showing off their ignorance. (Excluded from analysis)
2. Women should have to promise to obey their husbands when they get married. (Excluded from analysis)
3. Our country desperately needs a mighty leader who will do what has to be done to destroy the radical new ways and sinfulness that are ruining us.
4. Gays and lesbians are just as healthy and moral as anybody else. (Reverse-Scored)
5. It is always better to trust the judgment of the proper authorities in government and religion than to listen to the noisy rabble-rousers in our society who are trying to create doubt in people’s minds
6. Atheists and others who have rebelled against the established religions are no doubt every bit as good and virtuous as those who attend church regularly. (Reverse-Scored)
7. The only way our country can get through the crisis ahead is to get back to our traditional values, put some tough leaders in power, and silence the troublemakers spreading bad ideas.
8. There is absolutely nothing wrong with nudist camps. (Reverse-Scored)
9. Our country needs free thinkers who have the courage to defy traditional ways, even if this upsets many people. (Reverse-Scored)

10. Our country will be destroyed someday if we do not smash the perversions eating away at our moral fiber and traditional beliefs.
11. Everyone should have their own lifestyle, religious beliefs, and sexual preferences, even if it makes them different from everyone else. (Reverse-Scored)
12. The “old-fashioned ways” and the “old-fashioned values” still show the best way to live.
13. You have to admire those who challenged the law and the majority’s view by protesting for women’s abortion rights, for animal rights, or to abolish school prayer. (Reverse-Scored)
14. What our country really needs is a strong, determined leader who will crush evil, and take us back to our true path.
15. Some of the best people in our country are those who are challenging our government, criticizing religion, and ignoring the “normal way things are supposed to be done.” (Reverse-Scored)
16. God’s laws about abortion, pornography and marriage must be strictly followed before it is too late, and those who break them must be strongly punished.
17. There are many radical, immoral people in our country today, who are trying to ruin it for their own godless purposes, whom the authorities should put out of action.
18. A “woman’s place” should be wherever she wants to be. The days when women are submissive to their husbands and social conventions belong strictly in the past. (Reverse-Scored)
19. Our country will be great if we honor the ways of our forefathers, do what the authorities tell us to do, and get rid of the “rotten apples” who are ruining everything.

20. There is no “ONE right way” to live life; everybody has to create their own way. (Reverse-Scored)
21. Homosexuals and feminists should be praised for being brave enough to defy “traditional family values. (Reverse-Scored)
22. This country would work a lot better if certain groups of troublemakers would just shut up and accept their group’s traditional place in society.

Appendix S

Abortion, Vaccine Position, and Political Orientation

A woman should be allowed to terminate her pregnancy if she believes it is the right thing to do.

- ○ Very Strongly Disagree (1)
- ○ Strongly Disagree (2)
- ○ Moderately Disagree (3)
- ○ Slightly Disagree (4)
- ○ Neither Agree nor Disagree (5)
- ○ Slightly Agree (6)
- ○ Moderately Agree (7)
- ○ Strongly Agree (8)
- ○ Very Strongly Agree (9)

Which of the statements below is closest to your own position on childhood vaccines?

- ○ Children should not receive any vaccines at all. (1)
- ○ Children should receive some, but not all recommended vaccines. (2)
- ○ Children should receive all recommended vaccines. (3)

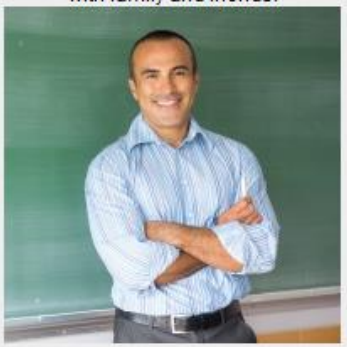
Politically, do you consider yourself:

- ○ Far-Left (1)
- ○ Liberal (2)
- ○ Moderately Liberal (3)
- ○ No Position/Center Moderate (4)
- ○ Moderately Conservative (5)
- ○ Conservative (6)
- ○ Far-Right (7)

Appendix T

Experiment 2 Character Images and Descriptions

Please read each description carefully, and tell us how much you think each person possesses each trait listed on the left side of the table below.

	Wesley earned his Masters in Health Services Administration at the University of Michigan. He is currently the executive vice president of Brattleboro Memorial Hospital in Vermont. He oversees the day-to-day operations of the hospital, including acting as a liason to the public, and planning for future projects and expansions to the hospital system. In his off time, he competes as an amateur billiard player and enjoys spending time with his 2 children and 3 dogs. 	Charles completed his medical degree at the University of New Mexico School of Medicine, did a residency at Tufts University, and a fellowship at the University of Chicago. He is board certified, and practices internal medicine at North Country Hospital in Newport, VT. When he is not working at the hospital, he spends Saturdays volunteering at a soup kitchen in his hometown, and enjoys hiking with his wife and 3 children. 	Adam earned his degree in Energy Healing at the Energy Healing School of Vibrational Medicine. He is the resident energy healer at Panacea - Healing for Body and Sole. He performs biofield therapy, reiki, and distance healing. In his off time, he serves on his local school board. 	Michael earned his Masters in Music Education at NYU. He is currently the music teacher at Foxborough High School in Massachusetts, teaching all band classes, including Jazz, Chorus, and Marching Band. He is an avid cyclist and enjoys cook-outs with family and friends. 

Appendix U

CAM Products and Services Included in Experiment 2 Questionnaire

Conventional treatments were included as fillers. They were excluded from analysis and are indicated with asterisks.

Products:

Please indicate which of the following health-related goods you have used in the past year.

Check all that apply.

Herbal medicines or supplements

Aspirin*

Multivitamins

Ibuprofen (Motrin, Advil)*

Homeopathic medicines (oscillococcinum, aconite, lycopodium, etc.)

Acetaminophen (Tylenol) *

Individual vitamin supplements

Topical pain relievers (Icy Hot, Bengay)*

Naproxen (Aleve)*

"Immune boost" supplements (Airborne, etc)

A self-detox kit or guide

Services:

Please indicate which of the following health-related services you have used in the past year. Check all that apply.

- Chiropractic
- Naturopathy
- Massage Therapy
- Dental Services, including oral surgery*
- Orthodontics*
- Physical Therapy*
- Reflexology
- Faith Healing
- Reiki or other Energy Healing
- General practice medical doctor*
- Acupuncture
- Osteopathy

Appendix V

Climate Change Acceptance Scale (Lewandowski, Oberauer, & Gignac, 2013)

Scale: 1 (Strongly Disagree) to 5 (Strongly Agree)

Please indicate your agreement or disagreement with the following statements about climate change.

1. I believe that burning fossil fuels increases atmospheric temperature to some measurable degree.
2. I believe that the burning of fossil fuels on the scale observed over the last 50 years has increased atmospheric temperature to an appreciable degree.
3. I believe that the burning of fossil fuels on the scale observed over the last 50 years will cause serious negative changes to the planet's climate unless there is a substantial switch to non-CO2-emitting energy sources.
4. I believe that the burning of fossil fuels on the scale observed over the last 50 years has caused serious negative changes to the planet's climate.

Appendix W

Strong arguments used in Study 3, with a seventh, brief opposing argument.

		Interrater Reliability κ
Introduction	On the next page, you will be asked to read and respond to a few arguments in favor of requiring University of Connecticut students to take a comprehensive exam in their major area as a requirement for graduation. The exam would be a test of what the student had learned after completing the major, and a certain score would be required if the student was to graduate. After reading each argument, you will have a short time to list your thoughts in response to the argument, and then report your degree of acceptance of the argument. Please click the >> button when you are ready to start. Please read each argument carefully.	n/a
1	"A study conducted by the Educational Testing Service revealed that most of the Ivy League Schools and several of the Big Ten Schools have senior comprehensive exams to maintain their academic excellence. Professors that were interviewed said that the	.72

	exams assure that only quality and knowledgeable students would be associated with the university. This increases the prestige of both the alumni and the academic institution."	
2	"Faculty members also revealed that the comprehensive exams would provide an incentive for students to study and to learn the course material rather than simply cramming to pass tests. A study conducted by the University of Notre Dame's School of Education revealed that universities with comprehensive exams have resisted the national trend of declining scores on standardized achievement tests. Average scores on achievement tests for the universities with comprehensive exams has risen over the last five years."	.73

3	"Graduate schools and law and medical schools are beginning to show a clear and significant preference for students who received their undergraduate degrees from institutions with comprehensive exams. As the Dean of the Stanford Law School recently wrote in the Chronicle of Higher Education - "Although Stanford has not, and will not discriminate on the basis of race, sex, or religion, we do show a strong preference for applicants who have demonstrated their expertise in an area of knowledge, and who have shown that they can master an area of study by passing a comprehensive exam at the undergraduate level." Thus, instituting a comprehensive exam will be an aid to those who seek admission to graduate and professional schools after graduation."	.81
4	"Students going on for post-bachelors study will not be the only ones helped, however. Faculty members at the University of Georgia have recently implemented comprehensive exams and they found that the average starting salary of their graduates increased $3,000 over a two-year period. At comparable schools without comprehensives, salaries increased by only $950. In interviews with employers, it was found that the comprehensive exam had attracted larger, more well-known corporations to recruit	.79

students. As Saul Siegel, a vice-president of Proctor and Gamble corporation was quoted as saying, "We are much less hesitant to offer the larger salaries and executive positions to these kids because by passing their area exam, they have guaranteed us that we are hiring an expert in the area rather than a person who may or may not be dependable and responsible. We are willing to pay a little more for that." The Educational Testing Service statistics confirmed that students who graduated from universities with comprehensive exams fared better on the job market than students who didn't have to take comprehensives and were placed into jobs reserved for those with masters degrees. Thus students who are going immediately on to the job market will benefit by the institution of comprehensive exams."

5 "An interesting feature of the comprehensive exam requirement is .76 that it has led to an improvement in the quality of undergraduate teaching in the schools where it has been instituted. Since departments look bad when their majors fail the test, more emphasis is placed on the quality of teaching and offering undergraduates a variety of important courses presented in an interesting and stimulating manner. Again, the Educational Testing Service statistics confirmed that teachers and courses at schools with comprehensive exams are rated more favorably by students than at schools with no exam. At the University of Florida, for instance,

teacher ratings increased significantly in the 3 year period after comprehensive exams were instituted."

6	"Finally, this university has recently suffered because the alumni have reduced their financial support even though inflation continues to rise. Increased alumni support is needed if quality programs are to be maintained, and continued tuition increases avoided. The chairman of the Board of Trustees has stated publicly that the alumni have refused to increase funding because of what they feel are lax educational standards at this institution. Alumni have assured the chairman of the University Board of Trustees that if academic standards and academic prestige were increased, more alumni contributions would follow. From all of the indications at schools which have adopted the comprehensive exam, the necessary increase in academic prestige and quality would accompany the institution of comprehensive exams. If fact, the prestigious National Accrediting Board of Higher Education which recently rejected this university's membership request, cited the lack of an undergraduate qualifying exam as one of its chief reasons for rejection. Accreditation by the NAB would enhance the university's reputation to graduate and professional schools, to employers, and demonstrate to alumni that the school is once again worth supporting. Thus, continued increases in tuition could be avoided."	.65

7	"Requiring all seniors to pass a comprehensive exam before they are allowed to graduate would damage the academic experience of University of Connecticut undergraduates and reduce their career prospects."	.69

κ scores represent interrater reliability of judging the number of favorable, unfavorable, and neutral statements for each argument.

Appendix X

Word Stem Procedure From Experiment 4

Instructions:

Now we have a short linguistic task for you. Please read the following instructions carefully.

In the next few pages, you will see 25 word stems - incomplete words that you will be asked to finish in as many ways as you can think of.

For example, if the word stem is "Sp___,"

You might type words like "spot," "spy," "spontaneous," etc. into the text field.

During this task, it is very important that you move quickly, so don't spend much time sitting still and thinking about which words would work. After 30 seconds, the page will automatically advance to the next word stem.

Let's try a practice run. Click the ">>" button at the bottom right of the screen to continue.

Separate words by starting a new line ("enter" or "return" key). Remember to go as fast as you can! Don't worry about small mistakes.

Ra____

Great! Now try another!

Remember to go as fast as you can. The page will automatically advance after 30 seconds.

To____

Now that you have a sense of how this task works, it is time to start the real trials.

Once again, remember to go as fast as you can. Don't worry about mistakes! Click the continue (>>) button as soon as you have to stop to think of new words that fit with the stem.

Click the >> button when you are ready to start.

The 25 stems are as follows. Each was presented on a separate webpage, displayed for 30 seconds before automatically progressing to the next.

La_____

St_____

Ba_____

Tr_____

Qu_____

Lo_____

Mo_____

Ga_____

Wo_____

Pa_____

Ne_____

Wa_____

Ce_____

Sh_____

Ca_____

Wi_____

Ro_____

Be_____

Te_____

Mi_____

Pr_____

Sa_____

Na_____

Ma_____

Th_____

Appendix Y

Statement Pairs from Experiment 5

Each statement pair is followed by the following questions:

Regardless of whether you believe it, Assume for a moment that the first statement is true. Given that, how possible or impossible is it that the second statement is also true?

Scale: -10 (Completely Impossible) to 10 (Completely Possible)

Regardless of how you answered the previous question, please explain in your own words what conditions would be necessary to make these statements completely compatible with each other. If the statements are entirely compatible already, please say so.

Instructions:

You may find the next few items a little unusual. For each item, you will be presented with two statements. For each of these pairs, you will be asked to do two things:

First, you will be asked to judge how possible the *second* statement is, IF the *first* statement is true.

Second, you will be asked to briefly write down what conditions would be necessary for both statements to be true at the same time. Please fill out this part of the questionnaire regardless of the possibility rating you give the pair. If the statements are already fully compatible, simply say so.

Statement Pairs

Pair 1
First Statement:
Richard Wright wrote *Native Son,* a novel of black life in Chicago, and *Black Boy*, which is highly autobiographical.

Second Statement:
Yang is the positive force in nature, as recognized by the Chinese religion.

Pair 2
First Statement:
Caffeine is harmful to the human brain.

Second Statement:
Caffeine enhances emotional health and intellectual ability.

Pair 3
First Statement:
The phrase, "government of the people, by the people, for the people" was used in the Gettysburg Address.

Second Statement:
10 degrees Celcius is equivalent to 50 degrees Fahrenheit.

Pair 4
First Statement:
Hector Drake is the fastest runner in human history.

Second Statement:
Sherri Harrison is the fastest runner in human history.

Pair 5
First Statement:
Vaccines prevent serious illness in children.

Second Statement:
Vaccines cause serious illness in children.

Pair 6
First Statement:
Faster-than-light travel is physically impossible.

Second Statement:
Some NASA craft have traveled 1 light-year in less than 11 months.

Table 1

T-tests for Student sample of Word Search pilot study

	Condition					
Dependent Variables	Distrust	Control	t	df	p	Cohen's d
Confidence	3.96	3.92	0.342	198	0.732	0.04
	0.84	0.813				
Control	3.95	3.85	0.918	198	0.360	0.13
	0.77	0.77				
IPIP Mistrust*	2.37	2.62	-2.564	199	0.011	0.12
	0.646	0.728				
Trust Neighbors	3.64	3.47	1.398	199	0.164	0.19
	0.855	0.904				
Trust Relatives*	4.24	3.92	2.507	198	0.013	0.35
	0.802	0.986				
Trust Self	4.41	4.2	1.735	194	0.084	0.24
	0.701	0.93				
Trust Science	3.66	3.7	-0.309	198	0.757	-0.04
	0.89	0.937				
Trust My Religion*	3.59	3.25	2.189	199	0.030	0.31

	1.06	1.167				
Yamagishi Trust	3.59	3.46	1.476	199	0.141	0.20
	0.619	0.648				
Yamagishi Caution	3.26	3.33	-1.007	198	0.315	-0.14
	0.494	0.508				
Routine Matchsticks	1.2	1.27	-0.398	215	0.691	-0.05
	1.316	1.332				
Non-Routine Matchsticks	0.15	0.18	-0.388	215	0.698	-0.05
	0.514	0.474				
Total Matchsticks	1.36	1.45	-0.467	215	0.641	-0.06
	1.525	1.554				
Similarity Judgments	3.84	3.81	0.23	197	0.818	0.03
	0.776	0.748				
IPIP Variety	3.34	3.35	-0.04	196	0.968	-0.01
	0.554	0.575				
IPIP Openness	3.45	3.39	0.873	194	0.384	0.12
	0.449	0.465				

Note. SDs are listed below means in the Condition columns.

Higher numbers represent a larger presence of that variable.

Table 2

T-tests For MTurk Sample of Word Search Pilot Study

	Condition						
Measures	Trust	Distrust	Control	F	df	p	Eta^2
Confidence	4.00	4.21	4.05	.35	2, 53	.707	.013
	.97	.71	.71				
Control	3.83	3.83	3.74	.06	2, 52	.940	.002
	.99	1.04	.81				
IPIP Mistrust	2.42	2.45	2.40	.01	2, 53	.987	.000
	.87	.86	.82				
Trust Neighbors	3.17	3.16	3.42	.45	2, 53	.640	.017
	.99	.83	1.07				
Trust Relatives	3.44	3.84	3.63	.59	2, 53	.556	.022
	1.29	.96	1.07				
Trust Self	4.22	4.32	4.16	.23	2, 53	.796	.009
	.65	.58	.90				
Trust Science	4.22	3.79	4.00	.99	2, 53	.378	.036
	.65	1.27	.75				
Trust My Religion	2.33	2.95	2.68	.92	2, 53	.406	.033
	1.28	1.39	1.46				

Yamagishi Trust	3.28	3.40	3.41	.12	2, 53	.884	.005
	1.01	.86	.88				
Yamagishi Caution	3.28	3.56	3.27	.95	2, 53	.392	.035
	.85	.71	.67				
Routine Matchsticks	1.83	1.47	1.95	.57	2, 53	.570	.021
	1.38	1.43	1.47				
Non-Routine Matchsticks	.06	.37	.58	1.99	2, 53	.146	.070
	.24	.83	1.07				
Total Matchsticks	1.89	1.84	2.53	.82	2, 53	.448	.030
	1.45	1.83	2.14				
Similarity Judgments	4.02	4.04	3.88	.23	2, 53	.797	.009
	.74	.69	.91				
IPIP Variety	3.40	3.23	3.40	.25	2, 53	.781	.009
	.89	.81	.89				
IPIP Openness	3.78	3.56	3.74	.58	2, 52	.577	.137
	.67	.61	.78				

Note. SDs are listed below means in the Condition columns.
Higher numbers mean more of the variable.

Table 3.

T-tests for Student sample of Scrambled Sentence Task pilot study

	Condition						
Dependent Variables	Trust	Distrust	Control	F	df	p	Eta^2
Confidence	3.90	3.83	4.08	1.41	110	.249	.025
	.68	.61	.67				
Control	3.90	3.89	3.95	.08	110	.922	.001
	.55	.79	.70				
IPIP Mistrust	2.24	2.38	2.45	1.09	110	.339	.019
	.57	.53	.79				
Trust Police	3.79	3.56	3.87	1.51	110	.225	.027
	.70	1.00	.704				
Trust Neighbors	3.82	3.69	3.58	1.02	110	.364	.018
	.76	.58	.86				
Trust Relatives*	$4.62^{\dagger}$	4.47	$4.26^{\dagger}$	3.35	110	.039	.057
	.54	.51	.72				
Trust Self	4.51	4.29	4.34	1.62	109	.203	.029
	.56	.62	.53				
Trust Science	3.95	3.72	4.05	2.10	110	.127	.037
	.69	.85	.57				
Trust My Religion	3.49	3.43	3.55	.13	109	.880	.002

	1.12	.95	1.06				
Yamagishi Trust	3.66	3.69	3.56	.62	110	.538	.011
	.54	.51	.54				
Yamagishi Caution	3.30	3.20	3.35	.88	109	.417	.016
	.45	.45	.50				
Routine Matchsticks	1.38	1.36	1.24	.16	110	.857	.003
	1.37	1.25	1.10				
Non-Routine Matchsticks	.21	.19	.32	.49	110	.614	.009
	.52	.58	.66				
Total Matchsticks	1.59	1.56	1.55	.01	110	.993	.000
	1.65	1.50	1.31				
Similarity Judgments	3.94	4.00	3.95	.07	110	.933	.001
	.76	.81	.77				
IPIP Variety	3.43	3.39	3.31	.46	110	.630	.008
	.57	.48	.57				
IPIP Openness	3.44	3.55	3.41	.90	109	.408	.016
	.50	.45	.41				

SDs are listed below means in the Condition columns.

*p<.05

†Groups differ significantly using the Bonferroni adjustment.

Table 4.

T-tests for MTurk sample of Scrambled Sentence Task pilot study

	Condition						
Dependent Variables	Trust	Distrust	Control	F	df	p	Eta^2
Confidence	4.24	4.12	4.24	.34	2, 52	.711	.013
	.44	.60	.44				
Control	3.88	3.76	3.81	.07	2, 52	.932	.003
	.70	.97	1.03				
IPIP Mistrust	2.32	2.18	2.23	.12	2, 52	.886	.005
	1.02	.73	.75				
Trust Police	3.41	3.53	2.95	1.39	2, 52	.258	.051
	.71	1.13	1.40				
Trust Neighbors	3.18	3.76	3.29	1.92	2, 52	.157	.069
	.95	.56	1.15				
Trust Relatives	3.94	3.82	3.15	2.09	2, 51	.135	.076
	1.39	.64	1.57				
Trust Self	4.41	4.18	4.48	.69	2, 52	.505	.026
	.80	.81	.81				
Trust Science	3.88	4.12	3.90	.39	2, 52	.680	.015
	.78	.93	.89				
Trust My Religion	3.18	3.65	3.38	.73	2, 52	.486	.027

	1.24	1.00	1.16				
Yamagishi Trust	3.39	3.69	3.37	.95	2, 52	.394	.035
	.79	.68	.91				
Yamagishi Caution	3.14	3.25	2.97	.62	2, 52	.544	.023
	.90	.64	.84				
Routine Matchsticks	1.76	1.06	1.86	2.20	2, 52	.121	.078
	1.35	1.20	1.20				
Non-Routine	.06	.12	.52	2.71	2, 52	.076	.094
Matchsticks							
	.24	.33	1.03				
Total Matchsticks*	1.82	1.18†	2.38†	3.23	2, 52	.048	.110
	1.29	1.19	1.75				
Similarity Judgments	4.29	3.96	4.24	1.30	2, 52	.282	.048
	.67	.56	.70				
IPIP Variety	3.55	3.05	3.51	2.32	2, 52	.109	.082
	.56	.80	.88				
IPIP Openness*	3.83	3.34†	3.87†	4.29	2, 52	.019	.142
	.56	.66	.57				

SDs are listed below means in the Condition columns.

*p<.05

†Groups differ significantly using the Bonferroni adjustment.

Table 5.

T-tests for Student sample of Impostor Detection pilot study

	Condition					
Dependent Variables	Distrust	Control	t	df	p	Cohen's d
Confidence*	3.39	3.77	-2.16	65.62	.035	-.51
	.84	.65				
Control	3.81	3.71	.41	69	.683	.11
	.89	.99				
IPIP Mistrust	2.25	2.34	-.53	62.42	.597	-.13
	.60	.81				
Trust Neighbors	3.78	3.63	.82	69	.414	.20
	.76	.77				
Trust Relatives*	4.53	4.14	2.05	69	.044	.49
	.70	.88				
Trust Self	4.28	4.23	.26	69	.795	06
	.74	.84				
Trust Science	3.83	3.69	.77	69	.442	.17
	.85	.76				
Trust My Religion	3.36	3.06	1.17	69	.245	.27
	1.05	1.14				
Yamagishi Trust	3.70	3.54	1.29	56.28	.202	.31

	.39	.63				
Yamagishi Caution	3.21	3.31	-.80	69	.429	-.19
	.43	.61				
Routine Matchsticks	1.25	1.63	-1.33	63.37	.189	-.32
	1.02	1.35				
Non-Routine Matchsticks	.28	.34	-.38	69	.702	-.08
	.70	.73				
Total Matchsticks	1.53	1.97	-1.24	69	.221	-.29
	1.32	1.69				
Similarity Judgments	3.75	3.97	-1.35	69	.182	-.32
	.76	.63				
IPIP Variety*	3.26	3.55	-2.33	69	.023	-.55
	.58	.47				
IPIP Openness	3.41	3.60	-1.54	68	.129	-.38
	.51	.50				

SDs are listed below means in the Condition columns.

*p<.05

Table 6.

T-tests for MTurk sample of Impostor Detection pilot study

	Condition					
Dependent Variables	Distrust	Control	t	df	p	Cohen's d
Confidence	3.80	3.84	-1.80	53	.846	-.05
	.76	.75				
Control	3.83	4.08	-1.17	53	.246	-.32
	.79	.76				
IPIP Mistrust*	2.55	2.00	2.49	53	.016	.67
	.83	.80				
Trust Neighbors	3.30	3.72	-1.8	53	.078	-.49
	.84	.89				
Trust Relatives*	3.77	4.36	-2.28	53	.027	-.62
	1.04	.86				
Trust Self	4.27	4.48	-1.17	53	.248	-.31
	.74	.59				
Trust Science	3.90	4.16	-1.13	53	.262	-.31
	.96	.69				
Trust My Religion	2.63	2.92	-.92	53	.364	-.25
	1.22	1.08				
Yamagishi Trust*	3.30	3.74	-2.05	53	.046	-.55

	.69	.90				
Yamagishi Caution	3.30	2.90	1.95	53	.056	.53
	.71	.80				
Routine Matchsticks	1.23	1.72	-1.36	53	.179	-.37
	1.14	1.51				
Non-Routine	.33	.36	-.13	53	.897	-.04
Matchsticks						
	.66	.86				
Total Matchsticks	1.57	2.08	-1.04	53	.302	-.28
	1.55	2.10				
Similarity Judgments	4.07	3.99	.56	49.06	.576	.14
	.69	.42				
IPIP Variety*	3.47	2.99	2.19	53	.033	.58
	.73	.91				
IPIP Openness	3.64	3.38	1.41	52	.165	.38
	.69	.68				

SDs are listed below means in the Condition columns.

*p<.05

Table 7.

Partner by Belief in partner 2x2 ANOVA tests for Student sample of Economic Deception Game pilot study

	Trustworthy Partner				Untrustworthy Partner				Interaction
Dependent Variables	Disbelieve in Partner	Believe in Partner	p	d	Dis-believe in Partner	Believe in Partner	p	d	p
Confidence	4.02	3.95	<.001	.08	3.15	3.15	.782	.00	.782
	.70	1.03			1.09	1.14			
Control	3.86	3.82	<.001	.05	3.41	3.33	.573	.08	.867
	.76	.893			.96	.96			
IPIP Mistrust	2.80	2.78	<.001	.06	2.94	2.98	.849	-.11	.469
	.32	.31			.33	.40			
Trust Neighbors	3.56	3.75	.400	-.23	3.47	3.67	.058	-.24	.991
	.91	.71			.90	.74			
Trust Relatives	4.26	4.51	.979	-.34	4.19	4.56	.001	-.53	.523

	.79	.66			.81	.55			
Trust Self	4.30	4.35	.300	-.07	4.35	4.49	.278	-.21	.634
	.68	.82			.67	.68			
Trust Science	3.99	4.02	.556	-.04	3.89	4.01	.479	-.16	.670
	.77	.80			.70	.79			
Trust My Religion	3.27	3.25	.205	.02	3.37	3.49	.718	-.11	.608
	1.13	1.03			1.05	1.17			
Yamagishi Trust	3.53	3.65	.663	-.21	3.56	3.56	.392	.00	.396
	.63	.51			.62	.52			
Yamagishi Caution	3.29	3.23	.845	.13	3.27	3.28	.677	-.02	.531
	.47	.44			.52	.43			
Routine Matchsticks	1.06	1.33	.012	-.22	.95	.74	.826	.20	.087
	1.11	1.32			1.06	.99			
Non-Routine	.11	.15	.295	-.09	.13	.00	.447	.33	.127
Matchsticks									
	.47	.47			.56	.00			

Total Matchsticks	1.17	1.48	.011	-.22	1.09	.74	.931	.30	.044
	1.30	1.48			1.32	.99			
Similarity Judgments	3.93	3.91	.505	.03	4.00	3.96	.712	.05	.850
	.66	.75			.74	.73			
IPIP Variety	3.51	3.44	.185	.12	3.38	3.37	.650	.02	.650
	.58	.63			.60	.69			
IPIP Openness	3.66	3.63	.018	.06	3.47	3.49	.672	-.03	.672
	.55	.49			.60	.56			

SDs are listed below means in the Condition columns.

Table 8.

Partner by Belief in partner 2x2 ANOVA tests for MTurk sample of Economic Deception Game pilot study

	Trustworthy Partner				Untrustworthy Partner				Interaction
Dependent Variables	Disbelieve in Partner	Believe in Partner	p	d	Dis-believe in Partner	Believe in Partner	p	d	p
Confidence	4.21	4.09	.643	.15	3.20	3.62	.021	.31	.394
	.80	.83			1.32	1.39			
Control	4.00	3.64	.604	.43	3.27	3.92	.430	-.57	.075
	.88	.81			1.10	1.19			
IPIP Mistrust	2.48	2.52	.228	-.08	2.90	2.46	.263	.68	.151
	.54	.46			.74	.54			
Trust Neighbors	3.57	3.73	.555	-.21	3.27	3.46	.338	-.15	.948
	.85	.65			1.22	1.33			
Trust Relatives	4.14	4.09	.327	.08	3.87	4.38	.970	-.50	.231

	.66	.54			.92	1.12			
Trust Self	4.43	4.36	.146	.12	4.20	4.77	.607	-.90	.069
	.51	.67			.78	.44			
Trust Science	3.93	4.18	.266	-.25	3.80	4.15	.773	-.37	.853
	.92	1.08			1.15	.69			
Trust My Religion	3.07	3.18	.846	-.08	3.13	3.15	.964	-.01	.906
	1.44	1.25			1.36	1.41			
Trust the Police	3.00	2.91	.748	-.07	3.07	3.38	.444	-.25	.563
	1.24	1.38			1.22	1.26			
Yamagishi Trust	3.58	3.76	.248	-.29	3.32	3.58	.459	-.29	.689
	.81	.37			1.00	.81			
Yamagishi Caution	3.23	3.28	.629	-.08	3.39	3.19	.850	.38	.442
	.79	.35			.58	.47			
Routine Matchsticks	1.14	1.45	.302	-.23	.67	1.08	.223	-.35	.887
	1.29	1.44			.90	1.38			
Non-Routine	.00	.27	.472	-.59	.13	.00	.472	.53	.040

Matchsticks									
	.00	.65			.35	.00			
Total Matchsticks	1.14	1.73	.249	-.39	.80	1.08	.185	-.23	.679
	1.29	1.68			1.01	1.38			
Similarity Judgments	4.27	4.14	.603	.17	4.21	4.11	.823	.14	.944
	.79	.73			.84	.58			
IPIP Variety	3.26	3.40	.018	-.19	3.08	3.98	.346	-1.14	.075
	.78	.67			.77	.81			
IPIP Openness	3.61	3.59	.227	.03	3.59	4.05	.228	-.67	.183
	.57	.61			.62	.75			

SDs are listed below means in the Condition columns.

Table 9

Study 1, Statement type Set 1

Statement type	Judged Plausibility Mean (SD)
True	60.83 (14.07)
False	52.20 (14.48)

Table 10

Study 1, Set 2 Statements Descriptive Statistics

	Statement Plausibility
Statement type	Mean (SD)
Irrelevant	58.57 (17.20)
Agreeing	59.62 (13.25)
True	66.77 (17.86)
False	48.89 (19.58)
Contradictory	49.46 (12.19)
True	61.95 (20.73)
False	41.13 (16.63)

Note. All irrelevant statements were false

Table 11

Study 1 multilevel models

Robust		Parameter	Null Model, Coefficient	Random Coefficients Model, Coefficient	Full Contextual Model, Coefficient	Model Excluding MMC, Coefficient
Fixed Effects/ Irrelevant Statements	Intercept	β_{00}	55.44***	58.60***	58.60***	58.60***
			(1.23)	(2.24)	(2.08)	(2.09)
	MMV Slope	β_{01}			10.60**	9.84**
					(3.96)	(3.03)
	MMC Slope	β_{02}			-1.37	
					(3.96)	
Agreeing Statements (T)	Intercept	β_{10}		8.49**	8.49**	8.49**
				(2.77)	(2.66)	(2.68)
	MMV Slope	β_{11}			-5.89	-8.58**
					(3.71)	(3.15)
	MMC Slope	β_{12}			-4.90	
					(5.01)	
Agreeing Statements (F)	Intercept	β_{20}		-9.78**	-9.78**	-9.78**
				(2.97)	(2.95)	(2.96)
	MMV Slope	β_{21}			-2.91	-3.96
					(4.31)	(3.74)

	MMC	β_{22}			-1.90	
	Slope				(5.89)	
Contradictory	Intercept	β_{30}		-18.21***	-18.21***	-18.21***
Statements (F)				(3.17)	(3.01)	(3.01)
	MMV	β_{31}			-12.66*	-11.79**
	Slope				(6.28)	(4.89)
	MMC	β_{32}			1.60	
	Slope				(6.03)	
Contradictory	Intercept	β_{40}		3.73	3.73	3.73
Statements (T)				(3.11)	(2.95)	(2.99)
	MMV	β_{41}			-14.49*	-10.42*
	Slope				(6.34)	(5.07)
	MMC	β_{42}			7.42	
	Slope				(5.65)	
Variance	$Var(e_{ij}) = \sigma^2$		436.63	320.81	304.08	311.03
	$Var(r_{0i}) = \tau_{00}$		2.78	25.95**	25.41**	24.46**
Goodness-of-fit	AIC		2682.97	2617.00	2622.42	2617.23
	BIC		2689.15	2631.42	2657.45	2641.96
	Deviance		2676.97	2603.00	2588.42	2593.23
	Parameters		3	7	17	12

Note. Standard Errors are reported in parentheses for each parameter.

Table 12

Study 2 – Summary of medically-relevant trait ratings

Partner	Presentation:	Admin	Doctor	Quack	Teacher	Row Mean
Trustworthy	Alone	4.19 (.50)	4.84 (.32)	3.18 (.72)	3.75 (.52)	3.99 (.35)
	Together	3.85 (.56)	4.61 (.56)	3.44 (.85)	3.49 (.51)	3.85 (.46)
	Mean	3.95 (.56)	4.68 (.51)	3.36 (.81)	3.57 (.52)	3.89 (.43)
Untrustworthy	Alone	3.84 (.68)	4.78 (.33)	3.26 (1.01)	3.58 (.47)	3.86 (.50)
	Together	3.90 (.56)	4.71 (.44)	3.20 (.81)	3.48 (.51)	3.82 (.40)
	Mean	3.87 (.61)	4.74 (.40)	3.22 (.90)	3.52 (.50)	3.84 (.44)
	Character Mean	3.91 (.59)	4.71 (.46)	3.29 (.86)	3.54 (.51)	3.86 (.44)

Note. SDs are in parentheses below means

Table 13

Study 2: Summary of the distance from the midpoint of medically-relevant trait ratings

		Admin	Doctor	Quack	Teacher	Row Mean
Partner:	Presentation:					
Trustworthy	Alone	1.30 (.40)	1.84 (.32)	.72 (.49)	1.15 (.42)	1.25 (.27)
	Together	1.15 (.42)	1.66 (.47)	.92 (.49)	1.12 (.41)	1.21 (.33)
	Mean	1.20 (.41)	1.71 (.44)	.86 (.49)	1.13 (.42)	1.23 (.31)
Untrustworthy	Alone	1.17 (.43)	1.78 (.33)	.94 (.57)	1.17 (.41)	1.27 (.32)
	Together	1.16 (.44)	1.74 (.38)	.79 (.49)	1.20 (.43)	1.22 (.34)
	Mean	1.17 (.44)	1.76 (.36)	.85 (.52)	1.19 (.42)	1.24 (.33)
	Character Mean	1.18 (.42)	1.73 (.40)	.86 (.51)	1.16 (.42)	1.23 (.32)

Table 14

Study 3 – Mean Cognitive Responses to Arguments and Mean Agreement Rating

Pro-Exam Arguments	Favorable	Unfavorable	Neutral	Total	Rating
Argument 1	.52 (.89)	1.21 (1.21)	.52 (.89)	2.25 (1.29)	3.68 (1.71)
Argument 2	.93 (1.07)	.80 (1.00)	.16 (.48)	1.89 (1.11)	4.54 (1.50)
Argument 3	.54 (.79)	1.01 (1.02)	.18 (.45)	1.72 (1.01)	4.02 (1.68)
Argument 4	.73 (.88)	1.08 (1.14)	.12 (.36)	1.92 (1.13)	4.15 (1.63)
Argument 5	.94 (1.04)	.69 (1.09)	.19 (.53)	1.81 (1.11)	4.59 (1.74)
Argument 6	.54 (.82)	.88 (1.13)	.26 (.64)	1.69 (1.17)	4.19 (1.59)
Pro-Exam Mean	.70 (.61)	.94 (.73)	.24 (.32)	1.88 (.89)	4.19 (1.17)
Anti-Exam Argument					
Argument 7	.89 (1.24)	.81 (1.15)	.27 (.74)	1.97 (1.33)	3.97 (1.96)

Note. Standard deviations are in parentheses.

Table 15

Study 3 - Mean Response Numbers and Agreement Ratings for Arguments by Partner Type

	Favorable Thoughts		Unfavorable Thoughts		Neutral Thoughts		Total Thoughts		Agreement Ratings	
	Trust Condition	Distrust Condition ..	Trust	Distrust	Trust	Distrust	Trust	Distrust	Trust	Distrust
Pro-Exam Arguments:										
1	.56 (.87)	.49 (.91)	1.17 (1.16)	1.25 (1.27)	.46 (.87)	.58 (.91)	2.18 (1.33)	2.31 (1.25)	3.59 (1.76)	3.76 (1.66)
2	.85 (1.10)	1.01 (1.04)	.84 (.97)	.76 (1.03)	.11 (.45)	.21 (.50)	1.80 (1.04)	1.99 (1.16)	4.41 (1.59)	4.68 (1.40)
3	.52 (.85)	.55 (.73)	1.04 (.99)	.98 (1.06)	.18 (.47)	.19 (.42)	1.73 (1.01)	1.71 (1.01)	3.87 (1.71)	4.16 (1.65)
4	.63 (.89)	.83 (.87)	1.17 (1.20)	.99 (1.08)	.11 (.36)	.13 (.37)	1.91 (1.19)	1.94 (1.08)	4.05 (1.61)	4.25 (1.66)
5	.95 (1.08)	.93 (.99)	.76 (1.13)	.61 (1.04)	.18 (.47)	.20 (.58)	1.89 (1.18)	1.74 (1.04)	4.51 (1.75)	4.67 (1.73)
6	.67 (.94)	.41 (.65)	.81 (1.00)	.95 (1.24)	.23 (.64)	.30 (.64)	1.71 (1.23)	1.66 (1.10)	4.41 (1.46)	3.96 (1.69)

	Favorable		Unfavorable		Neutral		Total		Agreement	
	Trust	Distrust	Trust	Distrust	Trust	Distrust	Trust	Distrust	Trust	Distrust
Anti-Exam Argument:										
7	.76 (1.19)	1.03 (1.27)	.91 (1.21)	.71 (1.08)	.28 (.82)	.26 (.65)	1.95 (1.48)	2.00 (1.16)	4.08 (1.92)	3.87 (2.00)

Standard deviations are in parentheses.

Table 16

Study 4 – Mean completions per word stem

	Impostor Detection	
Responses	Control	Distrust
	6.12 (2.04)	5.95 (2.10)
Practice Trials		
1	5.40[a] (3.19)	4.70[a] (2.65)
2	6.14 (2.68)	5.95 (2.32)
"Real" Trials		
1	6.74 (2.66)	6.77 (2.72)
2	7.39 (2.92)	7.16 (2.70)
3	6.92 (2.71)	6.65 (2.74)
4	6.19 (2.61)	6.59 (2.70)
5	5.55 (2.55)	5.46 (2.42)
6	6.30 (2.76)	6.67 (2.95)
7	6.66 (2.53)	6.71 (2.87)
8	5.53 (2.38)	5.29 (2.71)
9	5.43 (2.64)	5.20 (2.64)
10	6.66 (2.70)	6.47 (3.16)
11	4.98 (2.40)	4.87 (2.22)
12	6.50 (2.76)	6.33 (2.79)
13	2.88 (1.59)	2.88 (1.72)

14	6.92 (3.77)	6.33 (2.99)
15	7.28 (3.06)	6.66 (3.19)
16	5.95 (2.69)	5.46 (2.52)
17	6.21 (2.85)	6.24 (3.15)
18	7.30 (2.76)	7.05 (2.91)
19	5.76 (2.51)	5.46 (2.56)
20	5.92 (2.69)	5.82 (2.87)
21	5.90 (2.60)	5.87 (2.30)
22	6.53 (2.82)	6.50 (2.99)
23	4.30 (2.23)	4.08 (2.25)
24	6.76 (2.76)	6.51 (2.58)
25	7.09 (3.01)	6.98 (3.14)

Note. Standard deviations are in parentheses.

[a] Means differ at $p < .05$

Table 17

Study 5 – Statement Responses by Pairing Type

Pairing Type	Words Per Response	Sentences Per Response	Judged Possibility
Unrelated	15.79[a] (10.94)	1.23 (.49)	3.83[a] (5.47)
Conflicting	19.15[b] (13.48)	1.29 (.60)	-2.73[b] (5.44)
Contradicting	16.70[a] (11.76)	1.22 (.51)	-3.59[b] (5.81)

Note. Standard deviations are in parentheses.

Different superscripts within the same response type indicates a significant difference.

Table 18

Study 5 – Statement Responses by Pairing and Trust Condition

Pairing Type and Condition	Number Words	Number Sentences	Judged Possibility
Unrelated			
Trust. Part.	15.78 (11.40)	1.22 (.47)	4.32 (5.38)
Untrust. Part.	15.91 (10.52)	1.24 (.51)	3.43 (5.53)
Conflicting			
Trust. Part.	19.62 (11.95)	1.28 (.58)	-2.57 (5.59)
Untrust. Part.	18.81 (14.92)	1.29 (.61)	-2.95 (5.34)
Contradicting			
Trust. Part.	17.21 (11.42)	1.27 (.53)	-3.55 (5.77)
Untrust. Part.	16.31 (12.14)	1.18 (.50)	-3.66 (5.90)

Note. Standard deviations are in parentheses. Averages are assessed per response, not per participant..

Table 19

Study 6: Models of agreement with different arguments, including both manipulated distrust and medical mistrust.

			Anti-Rotation Message				
	Parameter	Pro-Rotation Message	General Agreement	Profit Motive	Not Worth Money	Not Worth Personal Risk	Responsible Behavior
			--Coefficients (SE)--				
Constant	B_0	6.33	3.01**	3.39**	3.24*	5.71***	4.26***
		(.72)	(.97)	(1.09)	(1.02)	(1.13)	(1.08)
Untrustworthy	B_1	.03	-4.91**	-2.50	-.14	-.75	-.68
Partner		(1.33)	(1.79)	(2.02)	(1.88)	(2.07)	(2.00)
MMV	B_2	-.106	.364	.46	.11	-.26	.27
		(.226)	(.30)	(.34)	(.32)	(.35)	(.34)
MMVxUP	B_3	-.097	1.41*	.67	.08	.163	-.07
		(.44)	(.60)	(.67)	(.63)	(.69)	(.67)
	Model F	.47	7.02	2.93	.09	.31	2.49
	Adjusted R^2	.02	.22	.08	-.05	-.03	.065
	p	.70	<.001	.04	.97	.82	.068

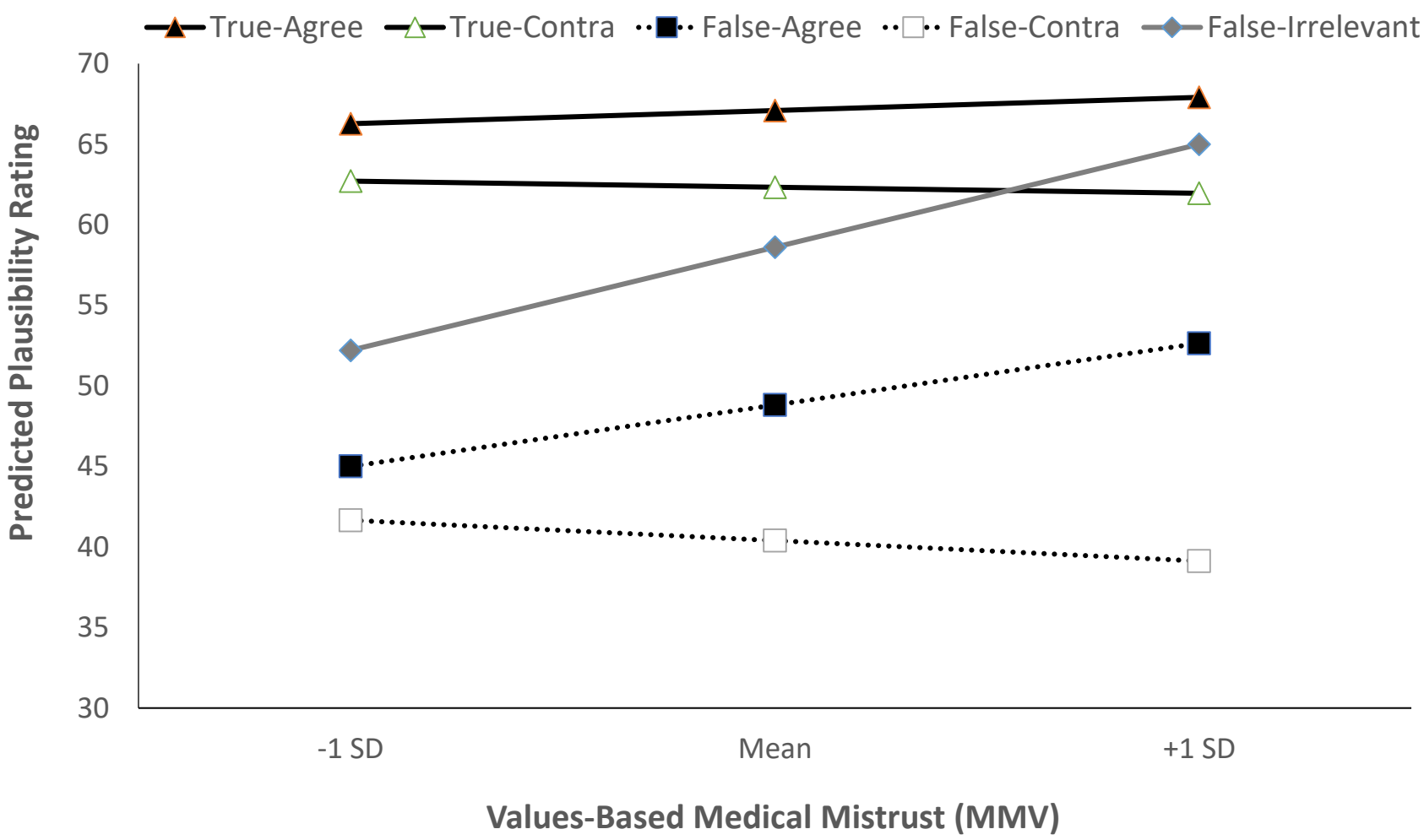


Figure 1. Estimated plausibility ratings based on MMV, statement pairing, and truth value of the statement, according to the multilevel model in Experiment 1.

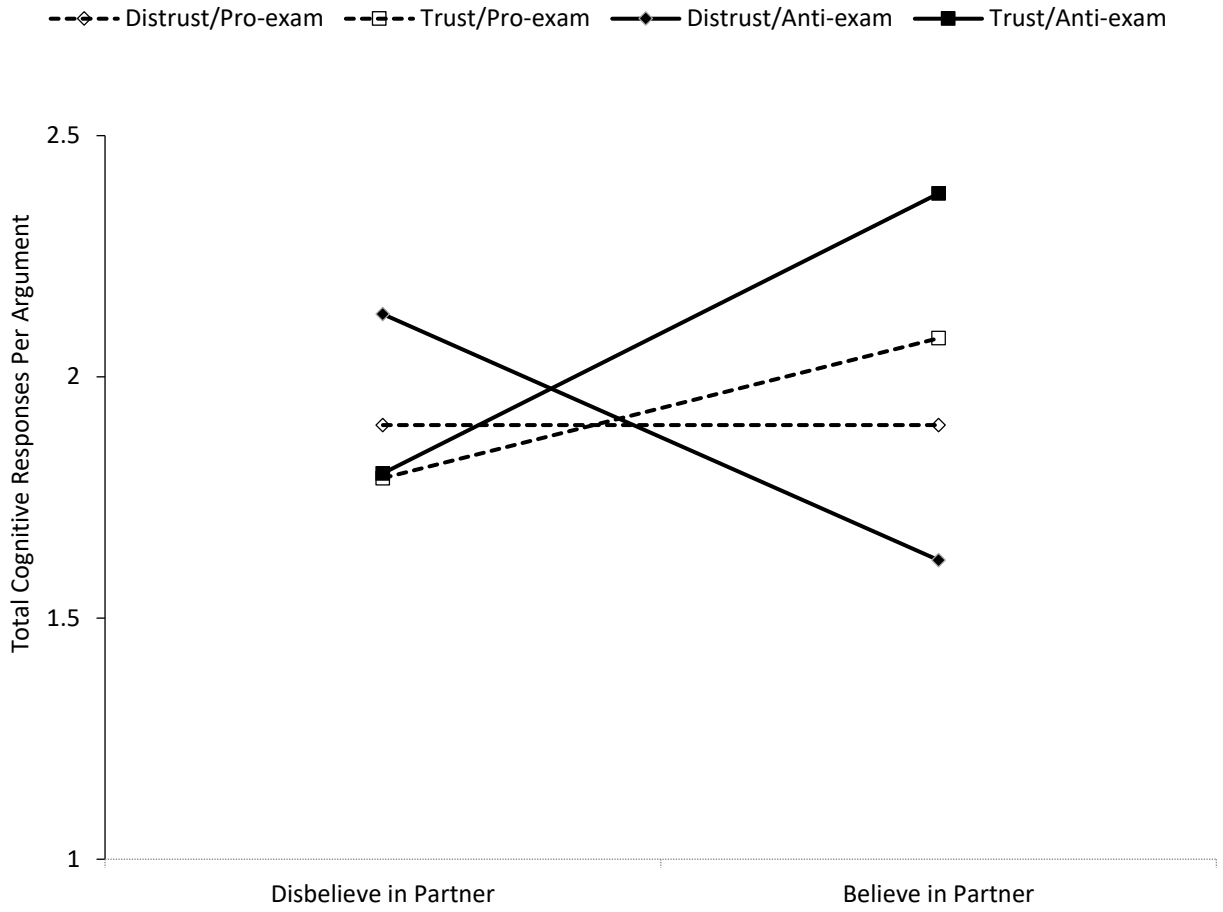


Figure 2. Interaction between belief in partner, trustworthiness of partner, and argument position on total cognitive responses.

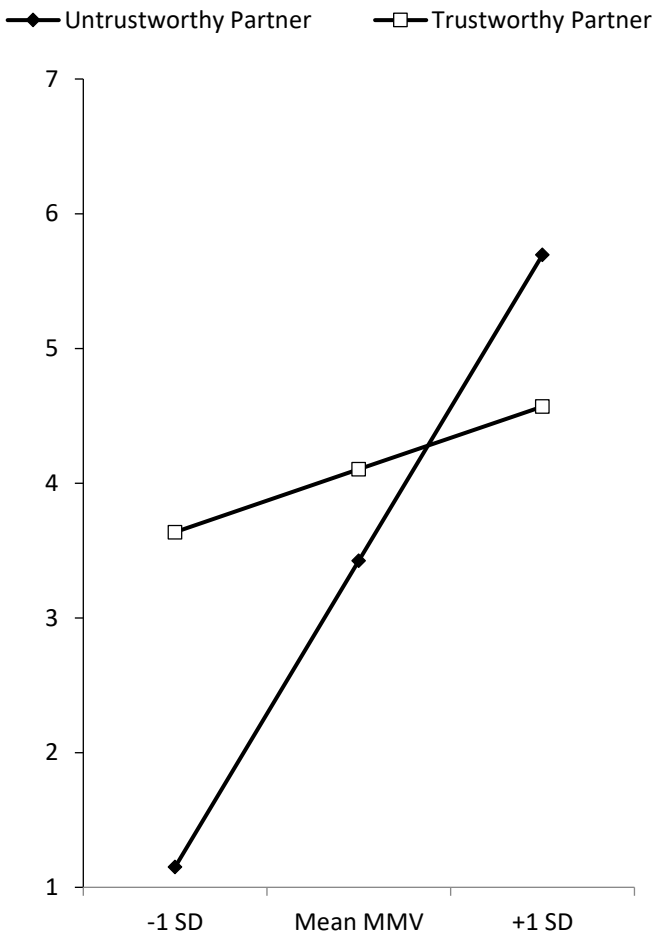


Figure 3. Estimated Overall Agreement With the Anti-Tire Rotation Essay in Experiment 6.